

UNITED STATES AIR FORCE
SUMMER RESEARCH PROGRAM -- 1996
SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 3B

PHILLIPS LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscas

Program Administrator, RDL
Johnetta Thompson

Program Administrator
Rebecca Kelly

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base
Washington, D.C.
December 1996

20010321 070

AGM01-06-1284

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering the data, reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Project, Washington, DC 20503.

ing and reviewing
for Information

0730

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December, 1996		3. F.	
4. TITLE AND SUBTITLE 1996 Summer Research Program (SRP), Summer Faculty Research Program (SFRP), Final Reports, Volume 3B, Phillips Laboratory				5. FUNDING NUMBERS F49620-93-C-0063	
6. AUTHOR(S) Gary Moore					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research & Development Laboratories (RDL) 5800 Uplander Way Culver City, CA 90230-6608				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) 801 N. Randolph St. Arlington, VA 22203-1977				10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES					
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the SFRP participants at the Phillips Laboratory.					
14. SUBJECT TERMS Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student				15. NUMBER OF PAGES	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL		

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 3 is bound in two parts, 3A and 3B. Volume 3A contains #1-22. Volume 3B contains reports #23-37. The Table of Contents for Volume 3 is included in both parts.

This document is one of a set of 16 volumes describing the 1996 AFOSR Summer Research Program. The following volumes comprise the set:

<u>VOLUME</u>	<u>TITLE</u>
1	Program Management Report
	<i>Summer Faculty Research Program (SFRP) Reports</i>
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A, 5B & 5C	Wright Laboratory
6	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	<i>Graduate Student Research Program (GSRP) Reports</i>
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center, and Wright Patterson Medical Center
	<i>High School Apprenticeship Program (HSAP) Reports</i>
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

SFRP FINAL REPORT TABLE OF CONTENTS

i-xii

1. INTRODUCTION	1
2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM	2
3. RECRUITING AND SELECTION	3
4. SITE VISITS	4
5. HBCU/MI PARTICIPATION	4
6. SRP FUNDING SOURCES	5
7. COMPENSATION FOR PARTICIPATIONS	5
8. CONTENTS OF THE 1996 REPORT	6

APPENDICIES:

A. PROGRAM STATISTICAL SUMMARY	A-1
B. SRP EVALUATION RESPONSES	B-1

SFRP FINAL REPORTS

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Richelle M Allen-King	Washington State University , Pullman , WA Reduction Kinetics in a Batch Metallic Iron/Water System:Effect of Iron/Water Exposure	AL/EQC	2- 1
DR Anthony R Andrews	Ohio University , Athens , OH Investigation of the Electrochemiluminescent Properties of Several Natural & Synthetic Compounds	AL/EQC	2- 2
DR Deborah L Armstrong	Univ of Texas at San Antonio , San Antonio , TX Development of A primary Cell Culture Preparation for Studying Mechanisms Governi ng Circadian Rhyth	AL/CFTO	2- 3
DR Robert L Armstrong	New Mexico State University , Las Cruces , NM Microparticle Bioluminescence	AL/CFD	2- 4
DR Maureen E Bronson	Wilkes Univ School of Pharmacy , Wilkes-Barre , PA Lack of Effect of UltraWideband Radiation on Pentylenetetrazol-Induced Convulsions in Rats	AL/OER	2- 5
DR Marc L Carter, PhD, PA	University of South Florida , Tampa , FL Assessment of the Reliability of Ground-Based Observers for the Detection of Aircraft	AL/OEO	2- 6
DR Jer-Sen Chen	Wright State University , Dayton , OH A Study of Data Compression Based on Human Visual Perception	AL/CFHV	2- 7
DR Cheng Cheng	Johns Hopkins University , Baltimore , MD Sequential Optimization Algorithm for Personnel Assignmt Based on Cut-Off Profiles & Rev of Brogden	AL/HRM	2- 8
DR Elizabeth T Davis	Georgia Institute of Tech , Atlanta , GA Perceptual Issues in Virtual Environments & Other Simulated Displays	AL/CFHP	2- 9
DR Keith F Eckerman	Univ of Tennessee , Knoxville , TN	AL/OEB	2- 10
DR Paul A Edwards	Edinboro Univ of Pennsylvania , Edinboro , PA A Viartion Fuel Identification- Neural Network Analysis of the Concentration of Benzene and Naphtha	AL/EQC	2- 11

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Randolph D Glickman	Univ of Texas Health Science Center , San Antonio , TX A Study of Oxidative Reactions Mediated by Laser-Excited Ocular Melanin	AL/OEO	2- 12
DR Ellen L Glickman-Weiss	Kent State University , Kent , OH The Effect of Short Duration Respiratory Musculature Training on Tactical Air Combat	AL/CFTF	2- 13
DR Irwin S Goldberg	St. Mary's Univ of San Antonio , San Antonio , TX Development of a Physiologically-Based Pharmacokinetic Model for the Uptake of Volatile Chemicals	AL/OES	2- 14
DR Robert J Hirko	University of Florida , Gainesville , FL Investigation of The Suitability of Tactile and Auditory Stimuli for use in Brain Actuated Control	AL/CFHP	2- 15
ISU VPP Acct4212313(Dooley)	Iowa State University , Ames , IA Determination of the Influence of Ultrawideband Exposure of Rats During Early Pregnancy on Pregnancy	AL/OER	2- 16
DR Andrew E Jackson	Arizona State University , Tempe , AZ A Description of Integrated Joint Use Initiatives to Satisfy Customer Requirements Across Govt Academia	AL/HRA	2- 17
DR John E Kalns	Ohio State University , Columbus , OH	AL/AOHR	2- 18
DR Nandini Kannan	Univ of Texas at San Antonio , San Antonio , TX Modeling Decompression Sickness Using Survival Analysis Techniques	AL/CFTS	2- 19
DR Antti J Koivo	Purdue Research Foundation , West Lafayette , IN Skill Evaluation of Human Operators	AL/CFBA	2- 20
DR Suk B Kong	Incarnate Word College , San Antonio , TX Aromatic Hydrocarbon Components in Diesel, Jet-A And JP-8 Fuels	AL/OEA	2- 21
DR Xuan Kong	Northern Illinois University , De Kalb , IL Mental Workload Classification via Physiological Signal Processing: EOG & EEG Analyses	AL/CFHP	2- 22

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Charles S Lessard	Texas A & M Univ-College Station , College Station , TX Preliminary Studies of Human Electroencephalogram (EEG) Correlates of GzAcceleration Tolerance	AL/CFTO	2- 23
DR Audrey D Levine	Utah State University , Logan , UT Biogeochemical Assessment of Natural Attenuation of JP-4 Contaminated Ground Water	AL/EQC	2- 24
DR David A Ludwig	Univ of N.C. at Greensboro , Greensboro , NC The Illusion of Control & Precision Associated w/Baseline Comparisons	AL/AOCY	2- 25
DR Robert G Main	Cal State Univ, Chico , Chico , CA Designing Instruction For Distance Learning	AL/HRT	2- 26
DR Phillip H Marshall	Texas Tech University , Lubbock , TX Time to Contact Judgments in The Presence of Static and Dynamic Objects: A Preliminary Report	AL/HRM	2- 27
MS Sandra L McAlister	Stonehill College , North Easton , MA	AL/AO	2- 28
MR Bruce V Mutter	Bluefield State College , Bluefield , WV Environmental Cost Analysis: Calculating Return on Investment for Emerging Technologies	AL/EQP	2- 29
DR Sundaram Narayanan	Wright State University , Dayton , OH Java-Based Application of the Model-View-Controller Framework in Developing Interfaces to interactive	AL/HRT	2- 30
DR Karl A Perusich	Purdue University , South Bend , IN Examining Alternate Entry Points in a Problem Using Fuzzy Cognitive Maps	AL/CFHI	2- 31
DR Judy L Ratliff	Murray State Univ , Murray , KY A Study of The Ability of Tunicates to be used as Global Bioindicators	AL/EQC	2- 32
DR Paul D Retzlaff	Univ of Northern Colorado , Greeley , CO Computerized Neuropsychological Assessment of USAF Pilots	AL/AOCN	2- 33

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR William G Rixey	University of Houston , Houston , TX The use of Solid-Phase Microextraction (SPME) for the low level Detection of BTEX and PAHs In Aqueou	AL/EQC _____	2- 34
DR Ali M Sadegh	CUNY-City College , New York , NY Investigation of Neck Models for Predicting Human Tolerance to Accelerations	AL/CFBE _____	2- 35
DR Kandasamy Selvavel	Claflin College , Orangeburg , SC Truncated Bivariate Exponential Models	AL/AOEP _____	2- 36
DR Barth F Smets	University of Connecticut , Storrs , CT Biodegradation of 2-4-DNT and 2,6-DNT in Mixed Culture Aerobic Fluidized Bed Reactor and Chemostat	AL/EQC _____	2- 37
DR Mary Alice Smith	University of Georgia , Athens , GA A Study of Apoptosis During Limb Development	AL/OET _____	2- 38
DR Daniel P Smith	Utah State University , Logan , UT Bioremediation & its Effect on Toxicity	AL/EQW _____	2- 39
MR. Joseph M Stauffer	Indiana State University , Terre Haute , IN Joint Corrections for Correlation Coefficients	AL/HRMA _____	2- 40
DR William B Stavinoha	Univ of Texas Health Science Center , San Antonio , TX Studies to Identify Characterisctic Changes in the Urine Following Ingestion of Poppy seed	AL/AOT _____	2- 41
DR William A Stock	Arizona State University , Tempe , AZ Application of Meta-Analysis to Research on Pilot Training	AL/HRA _____	2- 42
DR Nancy J Stone	Creighton University , Omaha , NE Engagement, Involvement, and Self-Regualted Leaarnign Construct and Measurement Development to Asses	AL/HRT _____	2- 43
DR Brenda M Sugrue	Univ of Northern Colorado , Greeley , CO Aptitude-Attribute Interactions in Test Performance	AL/HRTI _____	2- 44

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Stephen A Truhon	Winston-Salem State University, Winston-Salem, NC Mechanical Specialties in the U.S. Air Force: Accession Quality & Selection Test Validity	AL/HRM	2 - 45
DR Mariusz Ziejewski	North Dakota State University, Fargo, ND Validation of the Deformable Neck Model for A +Gz Acceleration	AL/CFBV	2 - 46

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Graham R Allan	New Mexico Highlands University, Las Vegas, NM Temporal and Spatial Characterization of a Synchronously-Pumped Periodically-Poled Lithium Niobate Optical	PL/LIDN	3 - 1
DR Brian P Beecken	Bethel College, St. Paul, MN Testing of a Dual-Band Infrared Focal Plane Array & An Infrared Camera Sys	PL/VTRP	3 - 2
DR Mikhail S Belen'kii	Georgia Inst of Technology, Atlanta, GA Tilt Sensing Technique w/Small Aperture Beam & Related Physical Phenomena	PL/LIG	3 - 3
DR Asoke K Bhattacharyya	Lincoln University, Jefferson City, MO Part A: Effect of Earth's Surface & Loss on the Resonant Frequencies of Buried Objects	PL/WSQ	3 - 4
DR Joseph M Calo	Brown University, Providence, RI Transient Studies of the Effects of Fire Suppressants in a Well-Stirred Combustor	PL/GPID	3 - 5
DR James J Carroll	Youngstown State University, Youngstown, OH Examination of Critical Issues in the use of (178) hf For High Energy Density Applications	PL/WSQ	3 - 6
DR Soyoung S Cha	Univ of Illinois at Chicago, Chicago, IL A Study on Hartmann Sensor Application to Flow Aero-Optics Investigation Through Tomographic Recons	PL/LIMS	3 - 7
DR Tsuchin Chu	Southern Illinois Univ-Carbondale, Carbondale, IL	PL/RKS	3 - 8
DR Kenneth Davies	Univ of Colorado at Boulder, Boulder, CO Studies of Ionospheric Electron contents and High-Frequency Radio Propagation	PL/GPIM	3 - 9
DR Judith E Dayhoff	Univ of Maryland, College Park, MD Dynamic Neural Networks: Prediction of an Air Jet Flowfield	PL/LIMS	3 - 10
DR Ronald R DeLyser	University of Denver, Denver, CO Analysis of Complex Cavities Using the Finite Difference Time Domain Method	PL/WSTS	3 - 11
DR Andrew G Detwiler	S Dakota School of Mines/Tech, Rapid City, SD Evaluation of Engine-Related Factors Influencing Contrail Prediction	PL/GPAB	3 - 12
DR Itzhak Dotan	The Open University of Israel, Tel-Aviv Israel Studies of Ion-Molecule Reaction Rates at Very High Temperatures	PL/GPID	3 - 13

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Omar S Es-Said	Loyola Marymount University, Los Angeles, CA On the Matis Selection of Durable Coatings for Cryogenic Engineer Technology	PL/RKE	3 - 14
DR Jeffrey F Friedman	University of Puerto Rico, Mayaguez, PR Testing the Frozen Screen Model of Atmospheric Turbulence	PL/LIMI	3 - 15
DR John A Guthrie	University of Central Oklahoma, Edmond, OK Ultrawide-Band Microwave Effects Testing on an Electronic System	PL/WSMA	3 - 16
DR George W Hanson	Univ of Wisconsin - Milwaukee, WI A Volumetric Eigenmode Expansion Method for Dielectric Bodies	PL/WSQ	3 - 17
DR Mayer Humi	Worcester Polytechnic Inst., Worcester, MA Wavelets and Their Applications to the Analysis of Meteorological Data	PL/GPAA	3 - 18
DR Christopher H Jenkins	S Dakota School of Mines/Tec, Rapid City, SD Shape Control of An Inflated Thin Circular Disk	PL/VT	3 - 19
DR Dikshitulu K Kalluri	University of Lowell, Lowell, MA Electromagnetic Wave Transformation in a Two-Dimensional-Space-Varying and Time-Varying Magnetoplasma	PL/GPIA	3 - 20
DR Aravinda Kar	University of Central Florida, Orlando, FL Thick Section Cutting w/Chemical Oxygen-Iodine Laser & Scaling Laws	PL/LIDB	3 - 21
DR Spencer P Kuo	Polytechnic University, Farmingdale, NY Theory of Electron Acceleration by HF-Excited Langmuir Waves	PL/GPI	3 - 23
DR Andre Y Lee	Michigan State University, East Lansing, MI Characterization Methods for Adhesion Strength Between Polymers & Ceramics	PL/RKS	3 - 24
DR Bruce W Liby	Manhattan College, Riverdale, NY Acousto-Optic Retro-Modulator	PL/VTRA	3 - 25
DR Feng-Bao Lin	Polytechnic Inst of New York, Brooklyn, NY Structural Ballistic Risk Assessment-Fracture Modeling	PL/RKEM	3 - 26
DR M Arfin K Lodhi	Texas Tech University, Lubbock, TX Theory, Modeling & Analysis of AMTEC	PL/VTP	3 - 27

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Ronald A Madler	Embry-Riddle Aeronautical University, Prescott, AZ Estimating the Area of Artificial Space Debris	PL/WSAT	3 - 28
DR Carlos A Ordonez	University of North Texas, Denton, TX Boundary Conditions at A Plasma-Facing Surface	PL/WSQA	3 - 29
DR Michael J Pangia	Georgia Southwestern Coll, Americus, GA Further Analysis of Kilohertz Order Waves Associated with Electron Beam Operations on STS46	PL/GPSG	3 - 30
DR Ronald M Pickett	University of Lowell, Lowell, MA Temporal-Displacement Stereograms of the Ionosphere: An Exploration of Their Utility in the Analysis of Equatorial Emission Depletion Bands	PL/GPIA	3 - 31
DR Edgar Sanchez-Sinencio	Texas A&M Univ-College Station, College Station, TX Low Voltage Analog Circuit Design for Radiation Tolerance	PL/VTER	3 - 32
DR Joseph C Slater	Wright State University, Dayton, OH Smart Structure/Actuator Modeling 7 Design for the Integrated Ground Demonstration Lab	PL/VTI	3 - 33
DR Ashok Srivastava	Louisiana State University, Baton Rouge, LA Modeling of Total Dose Response of SOI N-MOSFETS for Low Power CMOS Circuits	PL/VTER	3 - 34
DR James M Stiles	University of Kansas, Lawrence, KS The Potential Applications of Super-Resolution & Array Processing to Space-Based Radars	PL/VTRA	3 - 35
DR Charles M Swenson	Utah State University, Logan, UT Balloon Launch Retromodulator Experiment	PL/VTRA	3 - 36
DR Miguel Velez-Reyes	University of Puerto Rico, Mayaguez, PR Regularization Methods for Linear and Nonlinear Retrieval Problems	PL/GPAS	3 - 37

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
DR A F Anwar	University of Connecticut, Storrs, CT A Study of Quantum Wells Formed in AlxGa1-xAsySb1-y/InzGa1-zAs/AlxGa1-AsySb1-y Heterostructures	RL/ER	4 - 1
DR Ercument Arvas	Syracuse University, Syracuse, NY An Assessment of the Current State of the Art of Stap from an Electromagnetics Point of View	RL/OCSS	4 - 2
DR Ahmed E Barbour	Georgia Southern University, Statesboro, GA Formal Verification Using ORA Larch/VHDL Theorem Prover	RL/ERDD	4 - 3
DR Milica Barjaktarovic	Wilkes University, Wilkes Barre, PA Formal Specification and Verification of Missi Architecture Using Spin	RL/C3AB	4 - 4
DR Daniel C Bukofzer	Cal State Univ, Fresno, Fresno, CA Performance Analysis & Simulation Results of Delay & Spread Spectrum Modulated Flip Wave-Signal Gene	RL/C3BA	4 - 5
DR Xuesheng Chen	Wheaton College, Norton, MA Optical and Non-Destructive Methods to Determine the Composition and Thickness of an INxGA1-xAS/INP	RL/ERX	4 - 6
DR Jun Chen	Rochester Inst of Technology, Rochester, NY A Study of Optoelectronic Feedback-Sustained Pulsation of Laser Diodes at 1300 nm & 780 nm	RL/OCPA	4 - 7
DR Everett E Crisman	Brown University, Providence, RI Evaluation of Semiconductor Configurations as Sources for Optically Induced Microwave Pulses	RL/ERAC	4 - 8
DR Digendra K Das	SUNYIT, Utica, NY Techniques for Determining of the Precision of Reliability Predictions and Assessments.	RL/ERSR	4 - 9
DR Matthew E Edwards	Spelman College, Atlanta, Ga The Analysis of PROFILER for Modeling the Diffusion of Aluminum-Copper on a Silicon Substrate	RL/ERDR	4 - 10
DR Kaliappan Gopalan	Purdue University - Calumet, Hammond, IN Speaker Identification & Analysis of Stressed Speech	RL/IRAA	4 - 11
DR Joseph W Haus	Rensselaer Polytechnic Institute, Troy, NY Mode-Locked Laser Models and Simulations	RL/OCPA	4 - 12

SRP Final Report Table of Contents

Author	University/Institution	Rome Laboratory	Vol-Page
Report Title	Directorate		
DR James P LeBlanc		RL/OCSS	4 - 13
	New Mexico State University, Las Cruces, NM		
	Multichannel Autoregressive Modeling & Spectral Estimation Methods for Airborne Radar Environment		
DR David J McLaughlin		RL/ERCS	4 - 14
	Northeastern University, Boston, MA		
	A Review of Microwave Terrain Clutter Measurements at Bistatic		
DR Hrushikesh N Mhaskar		RL/ERAA	4 - 15
	Cal State Univ, Los Angeles, Los Angeles, Ca		
	Neural Beam Steering & Direction Finding		
DR Ronald W Noel		RL/C3CA	4 - 16
	Rensselaer Polytechnic Institute, Troy, NY		
	A Low Dimensional Categorization Technique for C Source Code		
DR Jeffrey B Norman		RL/OCPA	4 - 17
	Vassar College, Poughkeepsie, NY		
	Frequency Response of Semiconductor Photorefractive Matls: ZnTe:Mn:V,GaAs:Cr,&CdMnTe:V		
DR Glenn E Prescott		RL/C3BB	4 - 18
	University of Kansas Center for Research, Lawrence, KS		
	Rapid Prototyping of Software Radio Sys Using Field Programmable Gate Arrays & DSP Microprocessors		
DR Mark R Purtill		RL/C3CB	4 - 19
	Texas A&M Univ-Kingsville, Kingsville, TX		
	A Network Flow Heuristic for Graph Mapping		
DR Mysore R Rao		RL/OCSM	4 - 20
	Rochester Inst. Of Technology, Rochester, NY		
	Detection of Concealed Objects in Images: Investigation into Wavelet Transform Based Object Isolation Techniques		
DR Scott E Spetka		RL/IRD	4 - 21
	SUNY of Tech Utica, Utica, NY		
	Integrating a Multimedia Database & WWW Indexing Tools		
DR Gang Sun		RL/EROC	4 - 22
	University of Massachusetts-Boston, Boston, MA		
	Confined Optical Phonon Modes in Si/ZnS Superlattices		

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Mohammad S Alam	Purdue University, Fort Wayne, IN Fast Infrared Image Registration and High Resolution Reconstruction for Real Time Applications	WL/AAJT	5 - 1
DR Dominick Andrisani II	Purdue University, West Lafayette, IN A Fast Fourier Transform Analysis of Pilot Induced Oscillations	WL/FIGC	5 - 2
DR Pnina Ari-Gur	Western Michigan University, Kalamazoo, MI Texture and Microstructure of Hot Rolled Ti-6Al-4V	WL/MLLN	5 - 3
DR James D Baldwin	University of Oklahoma, Norman, OK Statistical Analysis of Fatigue Crack Growth Rate Data for 7075-T6 Aluminum Damaged by Prior Corrosion	WL/FIB	5 - 4
DR Armando R Barreto	Florida International Univ, Miami, FL Deconvolution of The Space-Time Radar Spectrum	WL/AAMR	5 - 5
MR Larry A Beardsley	Univ of Texas at Austin, Austin, TX The Use of Wavelets and Neural Networks in Data Compression, Data Fusion and Their Effects on Target Identification	WL/MNGA	5 - 6
DR Raj K Bhatnagar	University of Cincinnati, Cincinnati, OH Variable Width Template Construction for ATR with HRR Data	WL/AACR	5 - 7
DR Alley C Butler	University of Cincinnati, Cincinnati, OH Importance of Current Crowding and Self-Heating in a CdS/LaS Cold Cathode	WL/MLIM	5 - 9
DR Reaz A Chaudhuri	University of Utah, Salt Lake City, UT A Novel Compatibility/Equilibrium Based Iterative Post-Processing Approach for Axisymmetric Brittle	WL/MLBM	5 - 11
DR Julian Cheung	New York Inst. Of Technology, New York, NY New Techniques for Non-Cooperative Target Identification	WL/AACT	5 - 12
DR Milton Cone	Embry-Riddle Aeronautical University, Prescott, AZ Of Match Maker and Metrics	WL/AACF	5 - 13
DR Robert R Criss	Randolph-Macon Woman's College, Lynchburg, VA Optical Studies of Two Novel Electro-Explosive Devices	WL/MNMF	5 - 14

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Robert J DeAngelis	Univ of Nebraska - Lincoln, Lincoln, NE Granin Size Effects in the Determination of X-Ray Pole figures and Orientation Distribution Function	WL/MNM	5 - 15
DR Yujie J Ding	Bowling Green State University, Bowling Green, OH Investigation of Photoluminescence Intensity Saturation and Decay, and Nonlinear Optical Devices in Semiconductor Structures	WL/AADP	5 - 16
DR Gregory S Elliott	Rutgers State Univ of New Jersey, Piscataway, NJ Laser Based Diagnostic Techniques for Combustion and Compressible Flows	WL/POPT	5 - 17
DR Altan M Ferendeci	University of Cincinnati, Cincinnati, OH Vertical 3-D Interconnects for Multichip Modules	WL/AADI	5 - 18
DR Dennis R Flentge	Cedarville College, Cedarville, OH Kinetic Studies of the Thermal Decomposition of Demnum and X-1P Using the System for Thermal Diagnostic Studies (STDS)	WL/POSL	5 - 19
DR Himansu M Gajiwala	Tuskegee University, Tuskegee, AL Novel Approach for the Compressive Strength Improvement of Rigid Rod Polymers	WL/MLBP	5 - 20
DR Allen G Greenwood	Mississippi State University, Mississippi State, MS A Framework for Manufacturing-Oriented, Design-Directed Cost Estimation	WL/MTI	5 - 21
DR Rita A Gregory	Georgia Inst of Technology, Atlanta, GA Affects of Int'l Quality Standards on Bare Base Waste Disposal Alternatives	WL/FIVC	5 - 22
DR Michael A Grinfeld	Rutgers University, Piscataway, Piscataway, NJ Mismatch Stresses, Lamellar Microstructure & Mech	WL/MLLM	5 - 23
DR Awatef A Hamed	University of Cincinnati, Cincinnati, OH Inlet Distortion Test Considerations for High Cycle Fatigue in Gas Turbine Engines	WL/FIM	5 - 24
DR Stewart M Harris	SUNY Stony Brook, Stony Brook, NY Compositional Modulation During Epitaxial Growth of Some III-V Heterostructures	WL/MLPO	5 - 25
DR Larry S Helmick	Cedarville College, Cedarville, OH Effect of Humidity on Wear of M-50 Steel with a Krytox Lubricant	WL/MLBT	5 - 26
DR Kenneth L Hensley	University of Oklahoma, Norman, OK Hyperbaric Oxygen Effects on the Postischemic Brain	MED/SGP	5 - 27

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Iqbal Husain	University of Akron, Akron, OH Fault Analysis & Excitation Requirements for Switched Reluctance Starter-Generators	WL/POOC	5 - 28
DR David W Johnson	University of Dayton, Dayton, OH In Situ Formation of Standards for the Determination of Wear Metals in Perfluoropolyalkylether Lubricating Oils	WL/MLBT	5 - 29
DR Marian K Kazimierczuk	Wright State University, Dayton, OH Aircraft Super Capacitor Back-Up System	WL/POOC	5 - 30
DR Edward T Knobbe	Oklahoma State University, Stillwater, OK Corrosion Resistant Sol-Gel Coatings for Aircraft Aluminum Alloys	WL/MLBT	5 - 31
DR Michael C Larson	Tulane University, New Orleans, LA Cracks at Interfaces in Brittle Matrix Composites	WL/MLLM	5 - 32
DR Douglas A Lawrence	Ohio University, Athens, OH Analysis & Design of Gain Scheduled Missile Autopilots	WL/MNAG	5 - 33
DR Junghsen Lich	Wright State University, Dayton, OH Determination of 3D Deformations, Forces and Moments of Aircraft Tires with a Synchronized Optical and Analog System	WL/FIVM	5 - 34
DR Chun-Shin Lin	Univ of Missouri - Columbia, Columbia, MO Neural Network Technology for Pilot-Vehicle Interface & Decision Aids	WL/FIGP	5 - 35
DR Zongli Lin	SUNY Stony Brook, Stony Brook, NY Control of Linear Sys with Saturating Actuators with Applications to Flight Control Systems	WL/FI	5 - 36
DR Kuo-Chi Lin	University of Central Florida, Orlando, FL Study on Dead Reckoning Translation in High Level Architecture	WL/AASE	5 - 37
DR James S Marsh	University of West Florida, Pensacola, FL A Conceptual Model for Holographic Reconstruction & Minimizing Aberrations During Reconstruction	WL/MNSI	5 - 38
DR Paul Marshall	University of North Texas, Denton, TX Computational Studies of the Reactions of CH3I With H and OH	WL/MLBT	5 - 39

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Hui Meng	Kansas State University, Manhattan, KS Investigation of Holographic PIV and Holographic Visualization techniques for Fluid Flows and Flames	WL/POSC	5 - 40
DR Douglas J Miller	Cedarville College, Cedarville, OH Band Gap Calculations on Oligomers with an All-Carbon Backbone	WL/MLBP	5 - 41
DR Ravi K Nadella	Wilberforce University, Wilberforce, OH Hydrogen & Helium Ion Implantations for Obtaining High-Resistance Layers in N-Type 4H Silicon Carbide	WL/MLPO	5 - 42
DR Krishna Naishadham	Wright State University, Dayton, OH Hydrogen & Helium Ion Implantations for Obtaining High-Resistance Layers in N-Type 4H Silicon	WL/MLPO	5 - 43
DR Timothy S Newman	Univ of Alabama at Huntsville, Huntsville, AL A Summer Faculty Project for Anatomical Feature Extraction for Registration of Multiple Modalities of Brain MR	WL/AACR	5 - 44
DR Mohammed Y Niamat	University of Toledo, Toledo, OH FPGA Implementation of the Xpatch Ray Tracer	WL/AAST	5 - 45
DR James L Noyes	Wittenberg University, Springfield, OH The Development of New Learning Algorithms	WL/AACF	5 - 46
DR Anthony C Okafor	University of Missouri - Rolla, Rolla, MO Assessment of Developments in Machine Tool Technology	WL/MTI	5 - 47
DR Paul D Orkwis	University of Cincinnati, Cincinnati, OH Assessing the Suitability of the CFD++ Algorithm for Advanced Propulsion Concept simulations	WL/POPS	5 - 48
Dr Robert P Penno	University of Dayton, Dayton, OH Grating Lobes in Antenna Arrays	WL/AAMP	5 - 49
DR George A Petersson	Wesleyan University, Middletown, CT Absolute Rates for Chemical Reactions	WL/MLBT	5 - 50
DR Mohamed N Rahaman	University of Missouri - Rolla, Rolla, MO Effect of Solid Solution Additives on the Densification & Creep of Granular Ceramics	WL/MLLN	5 - 51

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Martin Schwartz	University of North Texas, Denton, TX AB Initio Modeling of the Enthalpies of Formation of Fluorocarbons	WL/MLBT	5 - 52
DR Thomas E Skinner	Wright State University, Dayton, OH A Method for Studying Changes in Tissue Energetics Resulting from Hyperbaric Oxygen Therapy	MED/SGP	5 - 53
DR Marek Skowronski	Carnegie Melon University, Pittsburgh, PA Investigation of Structural Defects in 4H-SiC Wafers	WL/MLPO	5 - 54
DR Grant D Smith	Univ of Missouri - Columbia, Columbia, MO Theoretical Investigation of Phthalocyanin Dimers	WL/MLPJ	5 - 55
DR James A Snide	University of Dayton, Dayton, OH Aging Aircraft: Preliminary Investigation of Various Materials and Process Issues	WL/MLLP	5 - 56
DR Yong D Song	North Carolina A & T State University, Greensboro, NC Memory-Base Control Methodology with Application to EMRAAT Missile	WL/MNAG	5 - 57
DR Raghavan Srinivasan	Wright State University, Dayton, OH Microstructural Development During Hot Deformation	WL/MLIM	5 - 58
DR Janusz A Starzyk	Ohio University, Athens, OH Feature Selection for ATR Neural Network Approach	WL/AACA	5 - 59
DR Alfred G Striz	University of Oklahoma, Norman, OK On Multiobjective Function Optimization in Engineering Design	WL/FIB	5 - 60
DR Barney E Taylor	Miami Univ - Hamilton, Hamilton, OH Optical and Electro-Optical Studies of Polymers	WL/MLBP	5 - 61
DR Joseph W Tedesco	Auburn University, Auburn, AL Effects of Airblast Characteristics on Structural Response	WL/MNSA	5 - 62
DR Scott K Thomas	Wright State University, Dayton, OH The Effects of Curvature on the Performance of a Spirally-Grooved Copper-Ethanol Heat Pipe	WL/POOS	5 - 63
DR James P Thomas	University of Notre Dame, Notre Dame, IN Subcritical Crack Growth of Ti-6Al-4V Under Ripple Loading Conditions	WL/MLLN	5 - 64
DR Karen A Tomko	Wright State University, Dayton, OH Grid Level Parallelization of an Implicit Solution of the 3D Navier-Stokes Equations	WL/FIM	5 - 65

SRP Final Report Table of Contents

Author	University/Institution	Arnold Engineering Development Center	Vol-Page
Report Title	Directorate		
DR Saad A Ahmed		AEDC	6 - 1
	King Fahd Univ of Petroleum & Minerals, Saudi, Arabia		
	Turbulence Statistics & Energy Budget of a Turbulent Shear Layer		
DR Csaba A Biegl		AEDC	6 - 2
	Vanderbilt University, Nashville, TN		
	Turbine Engine Blade Vibration Analysis System		
DR Frank G Collins		AEDC	6 - 3
	Tennessee Univ Space Institute, Tullahoma, TN		
	Laser Vapor Screen Flow Visualization Technique		
DR Randolph S Peterson		AEDC	6 - 4
	The University of the South, Sewanee, TN		
DR Robert L Roach		AEDC	6 - 5
	Tennessee Univ Space Institute, Tullahoma, TN		
	A Process for Setting Up Computation of Swirling Flows in the AEDC H-3 Heater		

SRP Final Report Table of Contents

<u>Author</u>	<u>University/Institution</u> <u>Report Title</u>	<u>U.S. Air Force Academy</u> <u>Directorate</u>	<u>Vol-Page</u>
DR Ryoichi Kawai	Univ of Alabama at Birmingham, Birmingham, AL A Massively Parallel Ab Initio Molecular Dynamics Simulation of Polymers & Molten Salts	USAFA _____	6 - 6

SRP Final Report Table of Contents

Author	University/Institution Report Title	Air Logistic Centers Directorate	Vol-Page
DR Sandra A Ashford	University of Detroit Mercy, Detroit, MI Evaluation of Current Jet Engine Performance Parameters Archive, Retrieval and Diagnostic System	OCALC	6 - 7
MR Jeffrey M Bigelow	Oklahoma Christian Univ of Science & Art, Oklahoma City, OK Enhancing Tinker's Raster-to-Vector Capabilities	OCALC	6 - 8
DR K M George	Oklahoma State University, Stillwater, OK A Computer Model for Sustainability Ranking	OCALC	6 - 9
DR Jagath J Kaluarachichi	Utah State University, Logan, UT Optimal Groundwater Management Using Genetic Algorithm	OCALC	6 - 10

INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from \$20,000 to \$25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1996 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After

completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education*, *Winds of Change*, and *IEEE Spectrum*. Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1996 are shown in this table.

1996 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	572	188	39
(HBCU/MI)	(119)	(27)	(5)
GSRP	235	109	7
(HBCU/MI)	(18)	(7)	(1)
HSAP	474	138	8
TOTAL	1281	435	54

4. SITE VISITS

During June and July of 1996, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCU/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant

increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1996 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1996 SRP selected participants are shown here.

1996 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	85	123
USAF Laboratory Funds	37	19	15
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	10	5	0
TOTAL	188	109	138

SFRP - 150 were selected, but nine canceled too late to be replaced.

GSRP - 90 were selected, but five canceled too late to be replaced (10 allocations for the ALCs were withheld by AFOSR.)

HSAP - 125 were selected, but two canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1996 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996
Faculty Members	\$690	\$718	\$740	\$740	\$740	\$770
Graduate Student (Master's Degree)	\$425	\$442	\$455	\$455	\$455	\$470
Graduate Student (Bachelor's Degree)	\$365	\$380	\$391	\$391	\$391	\$400
High School Student (First Year)	\$200	\$200	\$200	\$200	\$200	\$200
High School Student (Subsequent Years)	\$240	\$240	\$240	\$240	\$240	\$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of \$50/day for faculty and \$40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and

GSRP associates, 65 % (194 out of 297) claimed travel reimbursements at an average round-trip cost of \$780.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-five percent (85 out of 188) of faculty associates took orientation trips at an average cost of \$444. By contrast, in 1993, 58 % of SFRP associates took orientation visits at an average cost of \$685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

In 1996, RDL implemented direct deposit as a payment option for SFRP and GSRP associates. There were some growing pains. Of the 128 associates who opted for direct deposit, 17 did not check to ensure that their financial institutions could support direct deposit (and they couldn't), and eight associates never did provide RDL with their banks' ABA number (direct deposit bank routing number), so only 103 associates actually participated in the direct deposit program. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1996 REPORT

The complete set of reports for the 1996 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1996 associates, as indicated below:

1996 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. and Puerto Rico. Selectees represent 44 states plus Puerto Rico.

GSRP - Applicants came from 44 states and Puerto Rico. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	188
GSRP	109
HSAP	138
TOTAL	435

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	1	185
Master's	4	48	52
Bachelor's	0	60	60
TOTAL	188	109	297

SFRP Academic Titles	
Assistant Professor	79
Associate Professor	59
Professor	42
Instructor	3
Chairman	0
Visiting Professor	1
Visiting Assoc. Prof.	0
Research Associate	4
TOTAL	188

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
<i>IEEE Spectrum</i>	2%	1%
<i>BIIHE</i>	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

- A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).
- B. Faculty Associates suggest higher stipends for SFRP associates.
- C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1996 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

Lab	Evals Recv'd	How Many Associates Would You Prefer To Get ? (% Response)											
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
		0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
FJSRL	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
 - a. Time available for evaluation of applications:
 - b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5.
 - a. Benefits of associate's work to laboratory:
 - b. Benefits of associate's work to Air Force:
6.
 - a. Enhancement of research qualifications for LFP and staff:
 - b. Enhancement of research qualifications for SFRP associate:
 - c. Enhancement of research qualifications for GSRP associate:
7.
 - a. Enhancement of knowledge for LFP and staff:
 - b. Enhancement of knowledge for SFRP associate:
 - c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10.
 - a. Your working relationship with SFRP:
 - b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
 b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Responses to above questions

	<i>AEDC</i>	<i>AL</i>	<i>FJSRL</i>	<i>PL</i>	<i>RL</i>	<i>WHMC</i>	<i>WL</i>
<i># Evals Recv'd</i>	0	7	1	14	5	0	46
<i>Question #</i>							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1996 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1996 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13.
 - a. Would you apply again?
 - b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
 - a. At Home:
 - b. With Friend:
 - c. On Local Economy:
 - d. Base Quarters:
17. Value of orientation visit:
 - a. Essential:
 - b. Convenient:
 - c. Not Worth Cost:
 - d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1996 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1996 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1996 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

**CHARACTERIZATION METHODS FOR ADHESION STRENGTH
BETWEEN POLYMERS AND CERAMICS**

**Andre Lee
Assistant Professor
Department of Materials Science and Mechanics
Michigan State University
E. Lansing, MI. 48824**

**Final Report for:
Summer Faculty Research Program
Research Development Laboratories**

**Sponsored by:
Air Force Office of Scientific Research
Bolling AFB, D.C.**

and

**Phillips Laboratory
Edwards AFB, CA.**

CHARACTERIZATION METHODS FOR ADHESION STRENGTH BETWEEN POLYMERS AND CERAMICS

**Andre Lee
Assistant Professor
Department of Materials Science and Mechanics
Michigan State University**

ABSTRACT

Methods in characterization of adhesion strength between soft material to rigid materials were presented. The advantages and disadvantages of various methods were discussed in terms of fracture mechanics. Double cantilever beam method and double lap shear test are to be used to characterized the adhesion strength between the clay reinforced Nylon-6 to a ceramic material. The amount of material needed for these mechanical testing is much lesser than the normal dumb-bell shaped tensile specimens. Thus, these mechanical methods can be adapted by the chemist at the Phillips Laboratory where current development of POSS containing polymers can be examined in a "real" application environments.

CHARACTERIZATION METHODS FOR ADHESION STRENGTH BETWEEN POLYMERS AND CERAMICS

Andre Lee

INTRODUCTION

In many current engineering applications, often it is necessary to take advantages in properties from different classes of materials which may involve combinations of polymers, ceramics, and metals. A simplest method to combine these is a layered structure. Depending on the particular application, the strength of adhesion between layers of different materials may be critical. The "proper" experimental methods used to evaluate the adhesion strength depend on many factors such as the mechanical properties of the substrates and the adhesives, physical dimensions of the substrates and the adhesives, modes of principal loading (possible modes of fracture), directions of loading, service environmental considerations, etc.. These factors should be carefully considered prior to selecting the experimental methods to be used and correct evaluation of the experimental data such that the results can be best used for the engineering application of interest.

The critical application of interest to the Air Force's Phillips Laboratory is in the area of modern rocket propulsion. The hybrid polymer applications at the laboratory is focused on the development of new classes of ablative materials to reduce erosion rates. Currently, there are considerable efforts in the development of hybrid polymers, a new class of inorganic-organic thermoplastic co-polymers. However, another important area of research is to better understanding in the adhesion of these "hybrid" polymers to the ceramics casting. In this report, we described various possible methods that can be used to evaluate interfacial adhesion strength between hybrid polymers and ceramics. Due to the difficulties in the processing ceramic substrates, and complex mechanical testing, the results of mechanical testing will be reported separately in the form of professional journal paper. In the report, we limited in the discussion of various possible experimental methods for evaluating adhesive strength, processing of test specimen, and the analysis.

THEORETICAL CONSIDERATIONS

Since we are basically limited in the adhesion to the ceramic materials, the structures are said to contain soft adhesives or soft joints. A soft material is defined as one having low tensile, shear moduli and low hardness; that is, it can undergo large deformations at small loading. Of more significance to material testings and characterization are three other features often associated with soft materials: *nonlinear behavior, energy dissipation due to viscoelastic behavior, and high Poisson's ratio.*

It is clear that these properties impose additional requirements on testing techniques that necessitate mapping the nature and magnitude of the stresses. For this reason, it is advantageous to test soft joints with a method that does not necessitate a detailed knowledge of stress distributions and stress singularities and is capable of coping with energy dissipation. This can be achieved by using the energy balance approach of fracture energy. Note that linear fracture mechanics, LFM, is invalid when material do not behave in a linear elastic manner away from the crack tip. There are special cases, however, where test can be designed for evaluation by LFM or by the energy balance approach, interchangeably.

The basic tenet of fracture mechanics is that the strength of structures is governed by the initiation and propagation of inner flaws. Flaws can be introduced by a damage mechanisms during service. With the help of fracture mechanics, the resistance of a metal to flaw propagation is characterized by a parameter known as the fracture energy, G_C , or alternatively, the fracture toughness, K_{IC} . Because these parameters are true materials properties, independent of the geometry of the crack body, they can be used to predict ultimate strength for a given flaw size or critical flaw size for a given loading. Fracture energy and fracture toughness are interrelated as long as the behavior of the material is linear elastic away from the crack region. Because G_C is derived from the **energy balance approach**, it is the parameter of choice here for characterizing the fracture resistance of soft joints.

The energy balance approach was first used by Griffith¹ and Orowan² for brittle materials. According to this description, flaw propagation will occur whenever the stored elastic energy released per unit area of virtual crack is sufficient to overcome the cleavage energy of the materials.

It may be written:

$$\frac{\partial W}{\partial A} - \frac{\partial U}{\partial A} \geq 2\gamma \quad (1)$$

where W is work done by the external force, U is stored elastic energy, A is crack area, and γ is the surface free energy.

In joints containing highly deformable materials, the above criterion must be modified for two reasons: first, some mechanical energy may be converted to heat as a result of viscoelastic dissipation within the adherents, and second, the energy required to cause crack growth is far greater than twice the surface free energy because of dissipative mechanisms in the crack tip region. Thus, the criterion for crack growth becomes:

$$\frac{\partial W}{\partial A} - \frac{\partial U}{\partial A} - \frac{\partial V}{\partial A} \geq G_c \quad (2)$$

where V is the energy dissipated from the crack tip and G_c is fracture energy. The left side of the expression in Equation 2 is the energy release rate, or the driving force for debond, and is often denoted G. (As the driving force reflecting the actual state of stress, while the material property G_c describing the separation energy.) The dissipation term is significant when a soft adherent undergoes large deformation. This situation is encountered either when a soft adherent is peeled away from a rigid substrate or in blister specimen with a thin elastomeric membrane. The dissipation term can be evaluated in several manners: (1) extracting the energy loss from the hysteresis loop of the stress-strain curves³⁻⁵, (2) evaluate the viscoelastic constitutive behavior of the soft material via creep or relaxation studies, and then V can be calculated by subtracting the calculated stored energy from the external work, and (3) A rigorous approach the J-contour integral method⁶.

When possible, it is convenient to use fracture specimens that behave in a linear elastic manner, so the term $\partial V/\partial A$ can be neglected. In such cases, the Equation 2 can be shown to be equivalent to⁷:

$$\frac{F^2}{2} \frac{\partial C}{\partial A} \geq G_c \quad (3)$$

where F is the external load and C is the compliance of the structure. This expression is valid whenever the soft material is highly constrained between rigid adherents, as in the case of double cantilever beam specimen (wedge test specimen), as shown in Figure 1 and the butt joint specimen, shown in Figure 2. When applicable, Equation 3 is more convenient than that in Equation 2, because the compliance is easy to obtained experimentally by monitoring the stiffness of the structure. If the joint structure behaves in a nonlinear, nonelastic manner, the expression in Equation 3 is not applicable and the rate of energy dissipation must be evaluated. Failure to do so will yield an apparent fracture energy of $G_c + \partial V/\partial A$, which is not a true measure of adhesive fracture toughness.

The fracture energy G_c contains a dissipative component, even in the most brittle adhesives⁷:

$$G_c = G_0 + \psi \quad (4)$$

where G_0 is the energy required for rupturing secondary and primary bonds (intrinsic fracture energy), and ψ is the energy dissipated in viscoelastic and plastic deformation. In spite of the dissipative contributions, G_c remains an intrinsic material property as long as the dissipative mechanisms are confined to the crack tip region. In any case, the dissipated energy ψ is a function of rate and temperature. Therefore, we can confine the value of G_c by perform experiments at different rates and temperatures.

TESTING METHODS

Peel Tests:

Peel tests are widely used for evaluating the relative strength of adhesives and for comparing surface treatments. They usually consist of peeling away a soft adherent from a rigid substrate at a constant angle. The peel angle can be varied to obtain various combinations of mode I and mode II loading. (Mode I is refer as tensile loading, while mode II is in-plane shear.)

To maintain a constant angle and constant curvature of the peeled strip, a floating roller is required. Alternatively, angle control may be achieved by installing a stage capable of translating the substrate in concert with the crosshead. A number of variations on the peel test have been reported. Peel specimens are usually tested by applying a given displacement rate and recording the peel force versus time. A proper mechanical analysis then converts the obtained data to a peel energy versus peel rate curve.

The analysis of peel specimens can be quite involved for several reasons: (1) the peeled material typically has nonlinear, viscoelastic behavior and sometimes experience plastic deformation, (2) the peeled strip often undergoes large stretching and bending, and (3) a change in peel rate may drastically alter the mix of fracture modes at the crack tip. Extreme caution should be used, therefore when converting peel loads to fracture energy.

Despite these complexities, if we are only varying the surface treatments on the rigid, ceramic substrates, the peel test is capable to provide many useful information. However, since we are in the development of inorganic-organic hybrid polymers, the peel test is not attractive.

Blister Tests

In its most general form, the blister specimen consists of a layer of soft material completely adhering to a rigid substrate, except over a central portion where loading is applied⁷⁻⁹. The most common blister geometries have a circular unbonded region at the center. When the unbonded

region is pressurized by injection of a compressed fluid, the soft adherent lifts off and forms a blister-like cavity. The thickness of the soft material ranges from very thin (membrane) to very thick (semi-infinite medium). This method is sometimes referred as the cylindrical peel test.

When a critical pressure, P_C , or equivalently, a critical energy-release rate, G_C , is attained, the disc-shape flaw propagates radially. If one layer is transparent, crack initiation and propagation may be monitored optically. For opaque materials, alternate techniques based on pressure monitoring have been proposed⁸. Caution must be used when testing large blister specimens. Catastrophic failure may occur at constant pressure because the energy-release rate increases rapidly with debonded radius. Therefore, it is recommended to avoid designs that store too much elastic energy prior to reaching G_C , and to limit the flow rate of the pressurizing fluid with a bleed valve.

For blister test, in general, thin membranes should be avoided because deformations rapidly become too large as the blister grows, resulting in combined bending and stretching deformation that is difficult to analyze. This problem may be alleviated by constraining the membrane with a rigid plate or reinforcing the membrane with an elastic backing.

Double Cantilever Beam Test

A basic schematic of the double cantilever beam (DCB) specimen is shown in Figure 1. A extensometer is used for measuring compliance. The specimen consists of a sandwiched layer of soft adhesive material bonded between more rigid adherents, in our case the rigid adherent is the ceramic material. Because the soft material is constrained between two more rigid substrates, large deformations and nonlinear, nonelastic behavior away from the crack tip are avoided. If the beam behave elastically, the energy-release rate may be obtained experimentally by measuring compliance, C , versus crack distance, a .

If a DCB specimen is subjected to a **constant opening load**, the crack growth tends to be unstable because the energy-release rate increases with square of crack distance. However,

when loaded with a **constant displacement**, the DCB produces an energy-release rate that decays with the fourth power of crack distance, leading to a rapid crack arrest. Such behavior or test is observed in the wedge test, where a constant displacement is maintained using a shim or wedge inserted at the extremity.

To avoid rapidly changing energy-release rates, the DCB specimen can be tapered in either thickness or width direction of the substrates. In this case, the compliance varies linearly with crack distance, which means the driving force for debonds is independent of crack length. Thus the crack growth becomes stable.

DCB specimens are usually loaded perpendicular to the crack plane with two equal force acting in opposite directions. This type of loading yields pure mode I at the crack tip.

Compared with the peel and blister specimens, the DCB geometry greatly reduced viscoelastic dissipation in the soft material for three reasons¹⁰: (1) the volume of soft material is minimal, (2) the strain levels are smaller, thereby minimizing nonlinear effects, and (3) the soft layer is highly constrained, forcing it to deform in the bulk mode, rather than the shear mode. (The bulk modulus of polymers tends to be substantially less dissipative than the shear¹¹.) In the ideal case of a sandwiched layer with zero compliance and elastic beams, simple solutions based on elementary beam theory are available^{12,13}. Since, we are proposing to sandwich soft, polymer adhesive in between very rigid and elastic ceramic adherents, the simple solution may be applicable.

Tensile and Shear Tests

The butt joint specimen and the lap-shear specimen are commonly used to characterize adhesive strength. Typically, no macroscopic flaws are introduced and a strength-of-material approach is used in the analysis: the applied load at failure is simply converted to an average tensile strength or average shear strength by dividing load by area of the bonded region. Although this

simple approach can be useful for qualitative studies, to obtained failure parameters are useless for design purposes. Stress analysis of the butt joint and the lap-shear specimen have demonstrated that the distribution of these stresses is highly non uniform. Because of the complex stress states, the load at failure can not be readily converted into strength-of-material type of criterion, such as the maximum principal strain or the maximum octahedral stress. Furthermore, because of the absence of a well-defined crack, the load at failure can not be converted to a fracture energy or a fracture toughness. Thus, in the conventional approach, the tensile strength or shear strength reflects the presence of unknown flaws whose size may vary among specimens. For this reason, it is suggested that a better way to use tensile or shear tests is to introduce known flaws and to use an analysis based on fracture mechanics.

HYBRID POLYMERS

The hybrid polymers that are currently being developed at the Phillips's laboratory by Dr. Lichtenhan's group is a polymerizable monomers containing polyhedral oligomeric silsesquioxane (POSS)¹⁴. This new class of polymers have shown promise to low cost automatable processing, improved system reliability, design flexibility, environmentally acceptable technology, and improved ablation performance. However, at this time, the laboratory has not yet produce large amount of polymers for a systematic evaluation of interfacial strength to ceramic materials. Therefore, it was necessary to replace with commercial grades of "layered" silicate reinforced Nylon-6 nanocomposite.

The layered silicate systems are of montmorillonite clay that were exfoliated with ω -amino acid and ϵ -caprolactam condensed into Nylon-6. Therefore, the Nylon polymer chains are chemically attached to the exfoliated clay layers. The geometry of clay is a flat layer, with physical dimensions of 10 Å in the thickness and 1000 Å in the both width directions. The mechanical

properties of these layered reinforced systems had been demonstrated to show significant improvement over the non-reinforced Nylon-6. The adhesion of this hybrid polymer system to the ceramic materials are to be examined.

MECHANICAL EXPERIMENTS

As mentioned above, the hybrid polymer is considered to be the "soft" material in compared to the ceramic. With various complexity in mechanical tests as discussed above, the test geometry that we will be using are the **Double Cantilever Beam Test** and the **Double Lap Shear Test**. The advantages of these two tests are such we will be able to use linear elastic fracture mechanics to analyze the DCB test and the lap shear test will give us the bulk properties of the polymer in the true structural application conditions. Since both tests required very little amount of polymeric materials, if it can be demonstrated to be a good test. The small amount of POSS-containing polymers that have been made by the Phillips laboratory's group can also be examined. Mechanical tests are to be performed at various temperatures ranging from 25 °C up to 100 °C using either screw-driven mechanical tester or servo-hydraulic mechanical tester equipped with an oven. The mechanical tests are currently under way at Michigan State University. The results of mechanical experiments will be submitted to Dr. Lichtenhan and to reviewed journal paper for publication when completed.

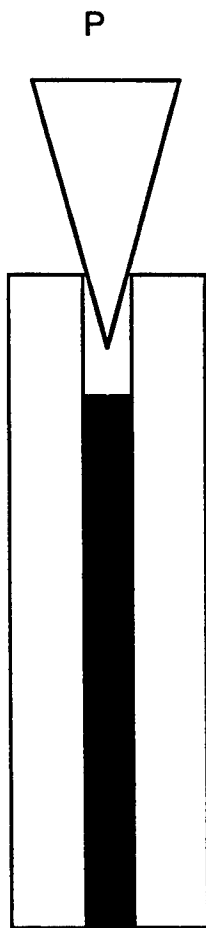
CONCLUSIONS

Mechanical properties characterizations are critically required for the continue advancement of hybrid polymeric materials. It is clear that these materials are to be used in a sandwich-like geometry to ceramic systems. In this report, we focused on various mechanical tests available for evaluate adhesion strength. Since the stiffness difference between polymer adhesives and ceramic

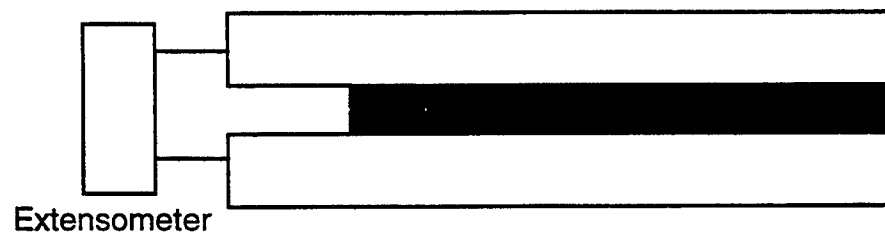
adherents is great, many linear elastic approximations can not be used. We have selected Double cantilever beam geometry, where by using this confined geometry, the nonlinear, nonelastic deformations is far away from the crack tip. Hence allows for energy balance approximation to be used in the data analysis. Furthermore, using the double lap shear testing, we will be able obtain bulk properties of polymer adhesive with small amount of polymers. This provided a great advantage to the chemists such that large variations of polymers can be evaluated and selected for the particular application of interest.

REFERENCES

1. A.A. Griffith, Philos. Trans. Soc. A221, 163 (1920).
2. E. Orowan, Rep. Prog. Phys. 12, 185, (1948).
3. A. Ahagon, A.N. Gent, H.J. Kim, and Y. Kumagi, Rubber Chem. Technol., 48, 896 (1975).
4. A. Kadir, and A.G. Thomas, J. of Polym. Sci. Polym. Phys. Ed., 22, 1623 (1984).
5. A.J. Kinloch, and A.D. Tod, Propel. Explos. Pyro, 9, 48 (1984).
6. A.J. Kinloch, and R.J. Younr, *Fracture Behavior of Polymers*, Applied Science Publisher, London (1983).
7. A. J. Kinloch, *Adhesion and Adhesives, Science and Technology*, Chapman and Hall, Cambridge (1987).
8. A.N. Gent and A.J. Kinloch, J. of Polym. Sci., Part A2, 9, 659 (1971).
9. E. H. Andrew, and A.J. Kinloch, Proc. R. Soc. A332, 385 (1973).
10. D.R. Lefebver, and D.A. Dillard, Exp. Mech. 28, 38 (1988).
11. J.D. Ferry, *Viscoelastic Properties of Polymerz*, 2nd edition, Wiley, New York, (1980).
12. G.P. Anderson. S.J. Bennett, and K.L. DeVries, *Analysis and Testing of Adhesive Bonds*, Academic Press, New York (1977).
13. A.N. Gent, Rubber chem. Technol., 47, 202 (1974).
14. J.D. Lichtenhan, *The Polymeric Materials Encyclopedia*, CRC Press, (1996).



(a)



(b)

Figure 1. (a) wedge geometry, where constant displacement is used to allow for crack arrest. (b) double cantilever beam geometry, where constant load is applied and extensometer measures the compliance. Often the crack growth is unstable.

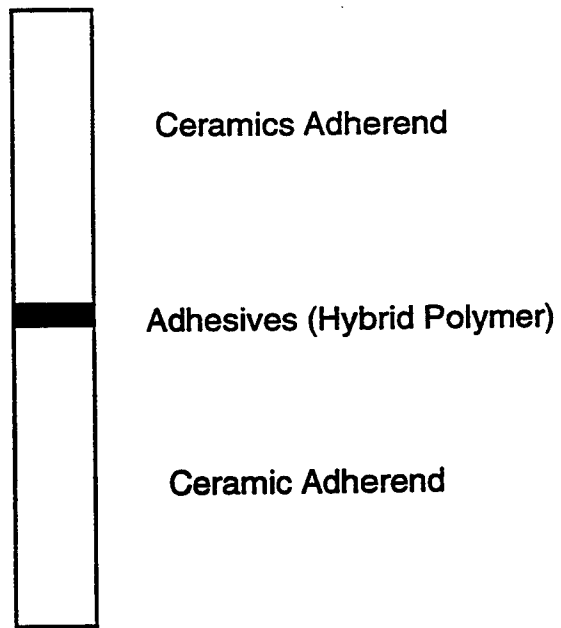


Figure 2. Butt joint specimen.

ACOUSTO-OPTIC RETRO-MODULATOR

Bruce W. Liby
Assistant Professor
Physics Department

Manhattan College
Riverdale, NY 10471

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling AFB, Washington, DC

and

Phillips Laboratory

August 1996

ACOUSTO-OPTIC RETRO-MODULATOR

Bruce W. Liby
Assistant Professor
Physics Department
Manhattan College
Riverdale, NY 10471

Abstract

In this paper a laser modulation device for space communications is proposed. It will utilize the acousto-optic effect in quartz to provide a rugged, monolithic, compact design for a space platform retro-modulator. In contrast with many other, modulation technique, the device can be polarization insensitive, able to tolerate a hostile environment, handle high laser intensities, and provide 1 - 1000 Mhz of modulation. All of this is based on currently available technology.

ACOUSTO-OPTIC RETRO-MODULATOR

Retro-modulation

A retro-modulator is a device that performs exactly as its name implies: it imposes a modulated signal on an input laser beam and retro-reflects that beam to the source¹. A detector co-located at the source receives the return beam. In this way a communications link is established. Retro-modulators offer significant advantages for a ground to space communications link. Much of the complex opto-electronics remain on the ground where it can be easily accessed and replaced, if needed. The weight required on the space platform is reduced, along with the power requirements. The key is to design a device that offers a high speed, low power modulator.

Many modulation techniques are commercially available and have been used for years in communications and engineering applications². The electro-optic effect is often exploited because of the maturity of its design, and because extremely high modulation bandwidths can be attained (>20 GHz). However, the size of such devices are limited because the voltage scales linearly with the aperture, thus beam sizes are limited to a few millimeters in most practical applications. Ferroelectric liquid crystals have large apertures with extremely low power requirements (<5 mW), but they are limited in speed to less than 50 kHz. Shutters add weight and mechanical complexity to the design. One attractive option is to use the acousto-optic effect in conjunction with a retro-reflector. Modulation bandwidths of greater than 10 Mhz are easily attained, and the devices have the added feature of being polarization insensitive.

Acousto-optic effect

In the acousto-optic effect, laser light is scattered off of sound waves in the acousto-optic medium (AOM). Many materials exhibit this behavior, including

such commonplace optics substrates like quartz. The sound wave is induced in the medium by attaching transducers to the crystal and driving them at optimum frequencies. This frequency is determined by velocity of sound in the medium and the wavelength of the laser light to be scattered. Devices such as simple deflectors and x-y scanners are commercially available. Acousto-optic modulators are also used as Q-switches in many laser systems.

The sound wave is produced by piezo-electric transducers attached to the crystal. The sound field modulates the index of the refraction of the field. The laser light can now scatter off of the index grating produced by the sound. Most acousto-optic modulators operate in what is referred to as the Bragg regime. This occurs when $L > \Delta^2/\lambda$, where L is the interaction length in the crystal, Δ is the wavelength of the sound in the crystal, and λ is the wavelength of light in the crystal. Light that is incident on the sound grating at the Bragg angle will be deflected at an angle of

$$\theta = \lambda * f / V$$

where θ is the deflection angle, f is the frequency of the sound wave and V is the speed of sound in the medium. The deflection angle can be increased by increasing the sound frequency within the bandwidth of the AOM. Deflection angles of 5 mrad are not difficult to attain.

The light can be scattered into higher orders, as in Bragg scattering. Care, however, must be taken to prevent the light from scattering back into the zeroth order which would result in no deflection. Typically acousto-optic modulators are designed to deflect into the first order. For this application any higher order is sufficient, thus decreasing the constraints on the system. The diffraction efficiency, η , is given by

$$\eta = \sin^2 \{ (\pi / \lambda) ([M_2 * P_{ac} * L] / [2 * H])^{1/2} \}$$

where M_2 is a "figure of merit," which is used as a yardstick to measure the acousto-optic material, P_{ac} is the acoustic power, L is the interaction length, and H is the optical beam width. By the proper choice of material and parameters, diffraction efficiencies of greater than 90% can be obtained.

All optically transparent materials are to some degree acousto-optic. Table 1 gives some pertinent parameters for a few materials⁴.

Material	Acoustic Velocity (km/sec)	Figure of Merit ($\times 10^{-15} \text{ m}^2/\text{W}$)
Flint Glass	3.51	8
Germanium	5.5	180
Lithium Niobate	6.6	7
Fused Quartz	5.96	1.6
Tellurium Oxide	5.5	1000

Table 1

A large figure of merit is not the only consideration. Some of the above materials require a particular wavelength polarization or may not be available in sufficient size. Cost and transparency must also be factored in.

The modulated signal will also undergo a frequency shift equal to the sound frequency. Thus precise frequency modulation can be achieved, if desired. This technique could conceivably be multiplexed with the beam deflection to increase the modulation bandwidth.

Geometrical considerations weigh heavily in the design of the modulator. Interaction length and beam waist affect the diffraction efficiency. Interference from backscattered sound waves can degrade the index grating. For a demonstration of retro-modulation this presents no real challenge; for a working device some sophistication must be used in the design.

Modulation via retro-reflection

The entire system consists of a CW laser source, detector and retro-modulator. The detector is mounted at the source. The laser illuminates the retro-modulator and the detector is, in turn, illuminated by the retro-modulator.

The main design feature is that the retro-reflector is also the modulator. This makes the unit compact and monolithic. Since many acousto-optic materials are optically transparent and easily fabricated, this should pose no real problem. Thus the difficulty associated with multiple optical elements is eliminated. The corner cube retro-reflector can even be silvered or HR coated to improve performance.

As mentioned previously, there are a number of ways to modulate the retro-reflected beam. In principle it is possible to modulate the phase, frequency or amplitude of the light. In this case we choose to deflect the beam to produce amplitude modulation, although amplitude and frequency modulation are directly available from the acousto-optic effect. When the beam, or a portion of it, is deflected, the intensity of the return will be reduced. The detector will read the change in intensity as amplitude modulation. Thus modulation will be achieved by simply spoiling the return beam, effectively turning it on and off. 100% modulation is not necessary; thus deflecting part of the beam, 10 to 20%, or perhaps even less, will be sufficient, depending on the sensitivity of the detector and diagnostics.

Thus the retro-modulator need only deflect the beam out of the detection area. For example, a retro-modulator on a space platform at an altitude of 30,000 m, that deflects the input signal 1 mrad will shift the beam at the source 30 m, well past the range of the detector. For shorter distances, the deflection will be correspondingly less. This presents no real problem, however, because the light can be spatially filtered to achieve the same effect.

Proposed device

A conceptual schematic of the proposed device is shown in Figure 1⁵. This proof-of-principle design is configured to provide a simple verification of the concept. Incoming light enters the retro-modulator through the vertical face, makes two reflections off of the back facets and is returned co-linearly to the source. The retro-modulator is essentially a corner cube retro-reflector that has been fabricated out of an acousto-optic material, such as flint glass. If the diameter of the beam fills the corner cube, then the light is simply inverted and reflected exactly back on itself.

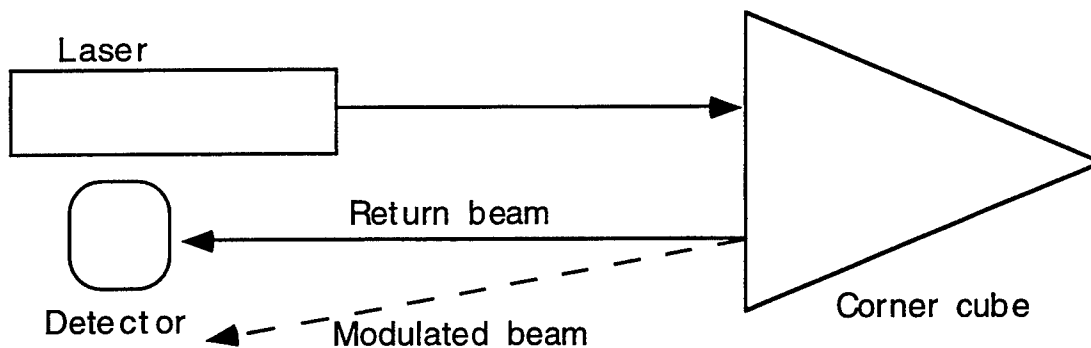


Figure 1

The transducers are attached to set up a sound field that will modulate the signal beam after its first reflection. They are positioned so that the part of the beam to be modulated is orthogonal to the induced grating. When the beam is modulated part of it will be reflected at an angle that spoils the retro-reflection.

The corner cube retro-reflector has three faces cut at 90° to each other. Three piezo-electric transducers are placed near the apex of the pyramid of the corner cube, as shown in Figure 2⁵. If each of them produces a sound wave of 3 mm diameter, they would cover a region 9 mm wide. If a 20 mm diameter beam is incident upon the corner cube, the transducers would create an index

grating that modulates 55% of the beam. If the diffraction efficiency is 85% (an arbitrary, yet attainable figure), then 47% of the beam would be deflected, and thus diverted, from the detector. This is, then, the modulation depth of the device. (In actuality it would be higher because the incoming beam would be subject to scattering into higher orders. The amount of this scattering is likely to be small, but any additional amount would only serve to increase the modulation depth.)

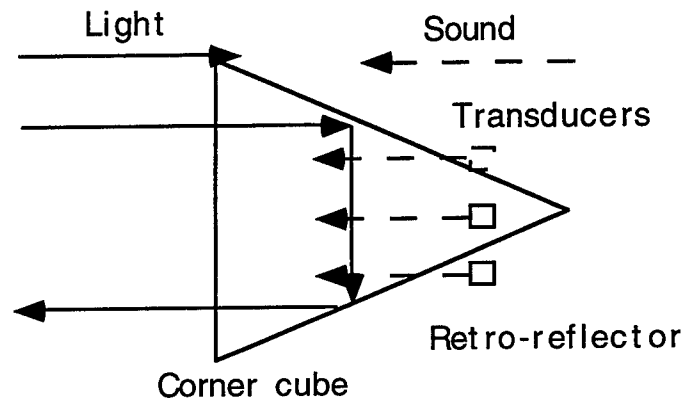


Figure 2

Care must be taken in the placement of the transducers to ensure good modulation. Parasitic effects, such as back reflections from the sound waves need to be considered before the design is finalized.

Suggestions for development

Once the device has been demonstrated as a viable technique, further testing and development would be required to evaluate its potential. Incident angle sensitivity would need to be determined and the unit optimized for an arbitrary input angle. A further refinement would be a Cat's eye retro-reflector. These devices bring the beam to a focus, reducing the diameter of the input beam. This factor could be exploited to increase the modulation even further than for the corner cube. Additional transducers could be added for further modulation. The capacity of the acousto-optic effect for amplitude modulation and

frequency modulation can be multiplexed with the beam deflection to increase the modulation bandwidth.

Summary

In this paper, a possible technique for an acousto-optic retro-modulator has been proposed. It uses the acousto-optic effect in a corner cube retro-reflector to deflect the reflected beam out of the detection area. It has the advantages of being monolithic, rugged, high speed (1 - 1000 MHz), low powered and simple in design. A demonstration experiment would be easy to carry out once the appropriate materials have been procured.

Acknowledgments

I would like to acknowledge the advice and suggestions of Dr. Charles Swenson of Utah State University, from whom I learned a great deal this past summer and Steve Dewalt, who shared his expertise of acousto-optics with me and reviewed this paper. I would also like to express my gratitude to my co-workers Clark Steed, also of Utah State University, and Teresa Hinojosa, Monica Obermier, and Fawn Miller, of the Phillips Lab, for their support and collegiality.

References

- 1 C. M. Swenson and G. L. Jensen, "Low-power transceiver for low earth orbit," to be published.
- 2 B. E. A. Saleh and M. C. Teich, "Fundamentals of Photonics," Wiley and Sons, Inc. New York, 1991.
- 3 J. Lekavich, "Basics of acousto-optic devices," Lasers and Applications, April, 1986.
- 4 Brimrose Technology Catalog, Baltimore, MD.
- 5 These figures are concept drawings, the angles and distances are not to scale.

STRUCTURAL BALLISTIC RISK ASSESSMENT – FRACTURE MODELING

Feng-Bao Lin
Associate Professor
Department of Civil and Environmental Engineering

Polytechnic University
6 Metrotech Center
Brooklyn, NY 11201

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, CA

and
Phillips Laboratory

August 1996

STRUCTURAL BALLISTIC RISK ASSESSMENT – FRACTURE MODELING

Feng-Bao Lin
Associate Professor
Department of Civil and Environmental Engineering
Polytechnic University

Abstract

The structural ballistic risk assessment of solid rocket motors requires a precise numerical simulation of the motors' firing. The simulation has to integrate technologies in various disciplines, including burnback, internal ballistics, and nonlinear structural analysis. One essential aspect in achieving a successful simulation is to consider the effect of cracks in propellant and debonds between propellant and insulator. The purpose of this research is to evaluate fracture mechanics models currently in the FEINT program and to identify new fracture models to be incorporated into FEINT so that its capability for predicting fracture behavior of motor firing can be improved.

In this report, the Thiokol's comprehensive study on the influence of defects on rocket motor firing performance is reviewed first. The fracture mechanics models currently in FEINT and their assumptions, computational approaches, and limitations are evaluated next. Several new fracture models such as Schapery's J_V integral, Hillerborg's fictitious crack model, Bazant's size-effect model, and Jenq and Shah's two-parameter fracture model are described, and their suitability to be applied to motor firing simulation is discussed. Finally, suggestions for achieving a better numerical simulation of crack and debond propagation in solid rocket motors are proposed.

STRUCTURAL BALLISTIC RISK ASSESSMENT – FRACTURE MODELING

Feng-Bao Lin

Introduction

While it is required to have a minimum of thirty years of visibility for solid rocket motor replacement, the current technology for motor structural integrity evaluation can only provide a visibility of about five years. This is due to the fact that the firing of a solid rocket motor itself is a complex process and its risk assessment requires an integration of advanced technologies in various disciplines. A typical solid rocket motor consists of case, insulator, liner (adhesive layer), and propellant. A numerical simulation of its firing involves burnback, internal ballistics, and structural analysis, in which one has to consider thermostrain, nonlinear constitutive relationship, large deformation, rate effect, interface modeling, debond and crack propagation, void growth, as well as the interactions between them. Among these, one essential aspect in the ballistic risk assessment is that cracks and voids in a propellant grain and debonds between propellant and insulator can exist in a solid rocket motor that could constitute a risk to its successful firing. These inherent anomalies could affect the ballistics performance of the motor, increase the bore pressure, and thus act as stress raisers and influence the motor's structural response. In order to predict, from a structural point of view, whether it is safe to fire a motor that contains those defects, fracture mechanics modeling would be a required component in the motor risk assessment system.

Between June 1988 and May 1995, Thiokol conducted a very thorough study on the ballistic risk assessment of solid rocket motors [Graham, etc., 1995]. They developed, integrated, and refined a system of computer software for predicting the performance of a solid rocket motor having natural or intentionally-made defects. This integrated computer programs, named Structural/Ballistic Analysis System (SBAS), include structural analysis codes: TEXLESP-2D and TEXLESP-3D [Collingwood, etc., 1988]; burnback analysis codes: RECESS and SURF; internal ballistics analysis codes: RECESS, ITTIB, and VOLFIL; crack/debond combustion codes: CCM and CDCA; and heat transfer codes: HEAT, ASCHAR, and SINDA. Furthermore, interfaces are provided for general commercial structural analysis codes: ABAQUS, ANSYS, NASTRAN, and TEXPAC. Any of these commercial codes can be selected to perform a structural analysis instead of using TEXLESP. The executive module for SBAS is the FEINT program. FEINT provides preprocessing (geometry definition, mesh generation, boundary conditions, material properties), analysis execution, and postprocessing (deformed geometry plots, contour plots, vector plots, XY plots). FEINT also provides interfaces between the various codes mentioned above. Thus, SBAS is often referred as FEINT. For preprocessing, the user can either build a complete finite element model within FEINT or import a model created in I-DEAS or PATRAN and modify it in FEINT.

One of the primary goals of Thiokol's study was to have a better understanding qualitatively and quantitatively of the influence of defects on the structural response of a solid rocket motor. This would provide a means of assessing the effect of motor defects on a static test firing and provide the test engineers with a rational

assessment of the risks of firing the motor so that actions can be taken to minimize damage to expensive test facilities.

Thirteen static firing tests of specially designed subscale motors were carried out in the Thiokol's program. To check the accuracy of the numerical results, Thiokol used the FEINT program to predict each motor performance before its firing test. The test measurements included pressure, strain, and deformation. The pressure measurements and the real-time radiography (RTR) observations together provided the best indication of the growth of defects. The predictions of the crack criticality were accurate, but debond and crack initiation and propagation speeds were generally overpredicted. It can be concluded that defects are critical to the performance of solid rocket motors, but considerable discrepancies between numerical predictions and test results exist.

Thus, the goal of this study, set forth by my summer research focal point, Durwood Thrasher, with my consent, is to identify advanced theoretical fracture mechanics models for near-term and far-term transition to practice for motor structural integrity assessment, and to define requirements for incorporating these models into the Structural Ballistic Analysis System, particularly focused on the finite element interface FEINT code. To achieve this goal, the tasks include:

1. Evaluate fracture mechanics models currently in FEINT, including assumptions, computational approaches, constitutive models, and limitations.
2. Assess ongoing research for applicable methodologies within Phillips Laboratory or from other resources.
3. Identify critical experiments for assessing or validating current models in FEINT, or for assessing advanced models to transition to practice.
4. Identify technical challenges of implementing new fracture models into FEINT, and approaches to overcome these challenges.

Fracture Mechanics Models in FEINT

The fracture module in FEINT currently consists of three different fracture analysis methods:

1. Crack closure integral method – This method calculates the energy release rate G based on the work required to reverse crack propagation by an amount ΔA :

$$G = \frac{dW}{dA} - \frac{dU}{dA} = \frac{\Delta W_c}{\Delta A} \quad (1)$$

The work required to reverse the crack, W_c , is the area under the force displacement curve. If the material behavior is such that the force-displacement path from the open crack state to the closed crack state is unique, as in a linear elastic material, then the crack closure integral may be applicable.

2. Crack opening displacement method – This method computes the stress intensity factors from the crack tip opening displacements [Anderson and Abrahamson, 1984], and is similar to that used by ANSYS. The method is applicable to cracks in isotropic linear elastic materials only. Based on the LEFM, the crack opening displacements Δv , Δu , and Δw (here Δ indicates displacement differences between two crack surfaces) near the crack tip are related to the stress intensity factors K_I , K_{II} , and K_{III} by:

$$\Delta v = \frac{(1+\kappa)K_I}{\mu\sqrt{2\pi}}\sqrt{r}, \quad \Delta u = \frac{(1+\kappa)K_{II}}{\mu\sqrt{2\pi}}\sqrt{r}, \quad \Delta w = \frac{4K_{III}}{\mu\sqrt{2\pi}}\sqrt{r} \quad (2)$$

where μ is the shear modulus, κ equals $(3 - \nu)/(1 + \nu)$ for plane stress problems and equals $(3 - 4\nu)$ for plane strain problems, and ν is the Poisson's ratio. Taking the natural log of both sides of the equations and solving for K as $r \rightarrow 0$ yields:

$$K_I = \frac{\mu\sqrt{2\pi}}{(1+\kappa)} e^{(\ln(\Delta v) - \frac{1}{2}\ln(r))}, \quad K_{II} = \frac{\mu\sqrt{2\pi}}{(1+\kappa)} e^{(\ln(\Delta u) - \frac{1}{2}\ln(r))}, \quad K_{III} = \frac{\mu\sqrt{2\pi}}{4} e^{(\ln(\Delta w) - \frac{1}{2}\ln(r))} \quad (3)$$

3. J-integral method – This method is based on the virtual crack extension method proposed by Parks [1974, 1977], and supports linear and nonlinear material behavior, including thermal strains and large deformation. The J-integral method implemented in FEINT also accounts for distributed loads on the crack surfaces. The procedure assumes that steady state conditions exist around the crack tip. This method can also calculate the energy release rate, stress intensity factors, and the speed of crack propagation if the necessary properties have been defined. Internodal constraints are ignored and should not be used inside the outermost contour of elements.

The J-integral is defined as the decrease in total potential energy, P , per unit area of crack extension, A , i.e.,

$$J = -\partial P / \partial A \quad (4)$$

Total potential energy is a function of material reference positions X , displacements u , and external loads f . For linear elastic materials, it can be shown that the J-integral is equal to the energy release rate G . For a virtual crack extension, δA , the above equation can be written in terms of virtual quantities as

$$J \delta A = -\delta P|_{u=\text{const.}} = (\delta W_E - \delta W_I)|_{u=\text{const.}} \quad (5)$$

where W_E and W_I are the work due to external and internal forces, respectively. Here, $u = \text{constant}$ means to keep the same displacements u during the virtual crack extension. This is due to the fact that the terms associated with δu vanish because the finite element solution yields the minimum potential energy, i.e., $\partial P / \partial u = 0$.

Internal work $W_I = \int_V U dV$, where U is the strain energy density and V is the entire volume of the solid under consideration. The variation of the internal work equation gives

$$\delta W_I = \int_V U \delta dV + \int_V \delta U dV \quad (6)$$

For linear elastic materials, $U = \frac{1}{2} \sigma \epsilon$ and $\delta U = \sigma \delta \epsilon$. For nonlinear materials, U depends on the material constitutive relationship and loading history, and is generally available as a step-by-step accumulation result from the analysis program. δU can be evaluated from a Taylor series expansion as $\delta U = \sigma \delta \epsilon + \frac{1}{2} \delta \sigma \delta \epsilon$. Since $\delta \sigma$ is unknown, the first order approximation $\delta U = \sigma \delta \epsilon$ is used, where Cauchy stress and the variation of the small strain induced by virtual crack extension are utilized in the FEINT program. It would be more accurate if the actual material model is used to determine δU .

Only loads in the vicinity of the crack will contribute to δW_E because a virtual crack extension produces displacement changes only in a small region near the crack tip. Restrict loads in the vicinity of the crack to normal pressure on the crack surface, and assume no change in load magnitude during the virtual crack extension. Then, we have

$$\delta W_E = \int_S p n \cdot u dS - \int_{S_0} p n \cdot u dS_0 \quad (7)$$

where p is the normal pressure applied to the current deformed crack surface, S_0 is the deformed crack surface, S is the deformed and extended crack surface, and n is the unit surface normal vector. Fig. 1 shows a virtual crack extension by perturbing the crack tip node. Note that the energy method does not allow the separation of K_I , K_{II} , and K_{III} in mixed mode situations, unless two of these fields are zero, i.e., a single mode fracture problem.

In FEINT, three crack tip mesh types, focused mesh, rectangular mesh, and zipper mesh, can be generated around a crack tip (Fig. 2). For a focused mesh, the user can specify the number of rings and the aspect ratio of elements in the area surrounding the crack tip. The default is that the length of the innermost ring of elements will be 1/50th of the total crack length and the element aspect ratio will be 2. For a rectangular mesh, the number of rings around the crack tip can be specified by the user, and the element size is reduced by a factor of two for each ring of elements toward the crack tip. Nodes on the boundary between rings of elements and not common to both rings are constrained to move with the outer ring of elements using constraint equations. The zipper crack tip mesh is particularly suitable for symmetry and bi-material crack mesh types. Ancillary nodes are generated between two crack surfaces so that the crack can be closed one element at a time from one step to the next. This is achieved by constrain the ancillary nodes through constraint equations in such a manner that when the ancillary

nodes are fixed, the crack is closed; when the ancillary nodes are free, the crack is free to open and the displacements of the ancillary nodes are the crack opening displacements between the crack surfaces.

In addition to the three fracture analysis methods, the crack/debond combustion codes, CCM [Derbidge, 1971] and CDCA [Lu, Yang, and Kuo, 1992], calculate the local pressure inside a crack. The analysis is critical in determining the magnitude of the loads acting on crack surfaces, and thus has a direct effect on crack propagation. On the other hand, cracks after propagation in a propellant expose additional surface area which can ignite and alter the pressure. Cracks with large burning-surface-area to exit-area ratios tend to have high pressures, which are caused by compressible choked-flow conditions. The CCM code computes a pressure distribution within a burning crack or debond by solving the governing equation for quasi-steady compressible flow with mass addition, assuming ideal gas law. The CDCA code simulates the transient crack combustion process. It considers pressure-wave propagation, ignition, flame spreading, and combustion process inside a crack or debond. The Noble-Abel dense-gas equation of state is used. Both codes so far have only one-dimensional simulation capability.

Fracture Predictions in the Subscale Motor Test Program

In the Thiokol's structural ballistic risk assessment program, they analyzed DMU motors, Minuteman motors, and subscale motors, and tested on Minuteman and subscale motors [Graham, etc., 1995]. The subscale motor program was particularly focused on the study of the effects of defects. This motor was designed to encourage defect propagation during firing. The motor case, made of filament wound nylon in a low-modules epoxy matrix, is highly flexible, and thus provides large strains in the propellant grain under ignition pressurization to allow crack or debond propagation. The case can reach 10 percent hoop strain prior to failure at about 900 psi of hydroburst test pressure. Meanwhile, catastrophic failure is not desired, so the motor is provided with thick internal insulation to protect the case from early flame exposure.

Numerical predictions using the FEINT program were made before each of thirteen static firing tests of subscale motors. The pressure measurements and the real-time radiography (RTR) observations provided the best indication of the growth of defects. A baseline motor without defects was first tested and test data were recorded. For numerical simulation, a coarse finite element model was used initially, allowing numerous analyses with different material properties and assumptions in order to match the test measurements as closely as possible. Then a fine finite element model was created and the best properties and analysis assumptions were used to predict the motor fracture behavior and to find fracture parameters. The flow charts of this analysis approach are shown in Fig. 3.

The results of large deformation structural analyses using ABAQUS based on the material properties determined in the coarse mesh were used to calculate the debond fracture energy at each step. A plot of debond fracture energy versus time was obtained. This debond fracture energy history was then used to calibrate a debond growth speed model using a modified power law, in which the crack speed is proportional to the difference between the current fracture energy and the critical energy, raised to a power:

$$da/dt = A (J - J_c)^q \quad (8)$$

In this model, both the critical fracture energy, J_c , and the crack speed coefficient, A , are rate dependent. This rate dependence is handled by making them another power law functions of the effective propellant modulus. The coefficients obtained by fitting the finite element results having an initial 0.1" deep debond, i.e., $a_0 = 0.1$ ", with the baseline motor test data were:

$$J_c = 0.00015 E_{eff}^{1.2}, \quad A = 1 \times 10^{-8} E_{eff}^{2.4}, \quad q = 0.83 \quad (9)$$

Unexpectedly, the fracture parameters fit to the static motor firing test data is considerably different from those obtained from laboratory specimen tests for a similar propellant material (Fig. 4). The critical fracture energy obtained from the static firing is lower than the laboratory data by an order of magnitude. Thiokol commented that it was not known whether the differences were from the material itself or because of the different configuration and loading between the motor and a laboratory specimen. However, if the critical fracture energy is a material property, it shall not be specimen geometry or loading configuration dependent. Clearly, more work is needed in identifying fracture properties of propellant. The fracture properties derived from the first motor static firing test were used in the subsequent numerical simulations.

Based on the test results and the failure analysis of the first motor, the motor was redesigned to include the addition of flaps at both ends of the grain and an increase in the nozzle throat diameter. Fig. 5 shows the flow chart for the subsequent analyses. Fracture energy calculated from the structural analysis was used to determine whether the debond or crack would propagate and how fast based on the debond speed model. The coupled burnback/ballistics analysis calculated the pressure history, which was applied as a load in the structural analysis. The structural analysis in turn provided strains as a function of the pressure, creating strain histories for computing time-dependent effective moduli for the propellant, the insulator, and the case.

Table 1 shows different defects and FEM models used in the subscale motor tests. Fig. 6 shows a 3-D finite element mesh for the tests with a half-circle plane bore crack. For the motors without defects, the predictions agreed well with the test data. But for those with cracks or debonds, they did not match well. Three predicted noncritical cracks and two predicted critical cracks were tested. The criticality of bore cracks was predicted correctly, but crack propagation speeds were much overpredicted. Debond propagation was predicted for all four of the axisymmetric debond motors, but was observed in none of them. A known natural debond in one motor was predicted to propagate; it did not. However, a similar debond in another motor propagated extensively. The differences between predictions and test measurements in test 10, 11, and 12 are shown in Fig. 7.

Fracture Mechanics Models Available for Solid Rocket Motors

A solid rocket propellant can in general be classified as a viscoelastic material. Its mechanical behavior is highly time, strain-rate, and temperature dependent. A good fracture mechanics model for a solid rocket propellant shall be able to capture these characteristics. The conventional J integral [Rice, 1968] will uniquely characterize the crack-tip conditions in a viscoelastic material only for a given failure time. This means a critical J value, J_{cr} , taken from a laboratory test is relevant to a structure only if they have similar failure times. Thus, it is difficult to apply such data to structural components. To account for the time dependence of viscoelastic materials, Schapery [1984] proposed a characterizing parameter, J_V , and showed that J_V uniquely defines the crack-tip conditions, making it a suitable fracture criterion for a wide range of rate-sensitive materials. By limiting the dependence of the time-shift factor on the stress or by choosing specific power-law forms for the constitutive response, Schapery [1990] further showed that J_V is path independent for nonlinearly viscoelastic materials with constitutive equations based on a reduced time, and is a candidate parameter for the characterization of the crack-tip fields. Bencher and his co-workers [1995] measured the fracture toughness of H-24 propellant at three different strain rates and three selected temperatures using Schapery's time-dependent viscoelastic approach. They concluded that the viscoelastic integral J_V appears to provide an appropriate measure of the toughness of propellant H-24.

A J-integral value can be computed from its elastic and plastic components from measurements of the applied load and load-point displacement using [Anderson, 1991]:

$$J = \frac{K_I^2}{E'} + \frac{\eta}{hb} \int_0^{\delta} P \cdot d\Delta \quad (10)$$

where E' equals E or $E/(1 - \nu^2)$ in plane stress or plane strain problems, respectively, h is the uncracked ligament length, b is the specimen thickness, P is the applied load, Δ is the load-point displacement, and η is a geometry-dependent factor.

The viscoelastic integral J_V can be calculated from the same formula by integrating a load vs. pseudoelastic displacement ($P - \Delta^e$) curve rather than a load vs. load-point displacement ($p - \Delta$) curve. The $P - \Delta^e$ curve is converted from a $p - \Delta$ curve by removing the creep displacements arising from the material's viscous behavior that occur during the fracture test. This was achieved by conducting a stress relaxation test at each of the relevant temperatures, and dividing the measured stress by the constant strain to calculate the effective Young's modulus as a function of time, $E(t)$. Then, the pseudoelastic displacement for a constant-displacement-rate fracture test is given by

$$\Delta^e = \frac{\bar{E}(t)}{E_R} \Delta \quad (11)$$

where reference modulus E_R is defined as the short-time relaxation modulus taken at time $t = 0$, and $\bar{E}(t)$ is the time-averaged Young's modulus given by

$$\bar{E}(t) = \frac{1}{t} \int_0^t E(t - \tau) d\tau \quad (12)$$

If a material is linear-viscoelastic, $\bar{E}(t)$ is independent of the strain applied in the stress relaxation test and the $P - \Delta^\circ$ curve after removing the relaxation strains will be linear.

Another critical aspect is the effect of inelastic fracture zone. The linear elastic stress analysis indicates that tensile stress at a sharp crack tip approaches infinity. Since material cannot withstand infinitely large stress, an inelastic zone must exist in front of the crack tip. In this inelastic zone, some fracture process, such as microcracking or microvoid formation, occurs. Thus, the zone is also termed as the fracture process zone. For homogeneous materials, because the size of the fracture process zone is negligible compared to the overall dimensions of the member examined, the effect of the presence of the zone is usually ignored, and thus linear elastic fracture mechanics (LEFM) can be approximately applied. However, this nonlinearity localized at the crack tip is usually more prominent in inhomogeneous materials since the fracture process zone spreads to a relatively large size due to the heterogeneity nature of the materials. Solid rocket propellants possess a microstructure dictated by agglomerates of filler particles embedded in polymer matrix, and are considered as inhomogeneous materials. In order to better simulate the fracture behavior of solid rocket motors, consideration of the fracture process zone effect is required. While LEFM requires only one fracture parameter to characterize the crack tip behavior, it requires more than one parameter to account for the influence of fracture process zone. Several models have been proposed:

1. Fictitious crack model – An early attempt to model the behavior in the fracture process zone was presented by Dugdale and is considered to be an extension of the linear elastic behavior to ductile materials. Because of its assumption of elastic perfectly plastic behavior, this model cannot accurately represent damage in inhomogeneous composites since it fails to account for factors such as microcracking and particle pullout. Hillerborg [1976] proposed a fictitious crack model, which is considered more suitable for inhomogeneous materials. The model is based on applying a cohesive stress, which decreases with the increase of the crack opening displacements rather than a constant value, on the crack surface near the crack tip. The fracture parameters of this model include the tensile strength, f_t , the crack opening displacement beyond which cohesive stress vanishes, w_c , and the energy release rate, G_F . G_F is defined as the area under the cohesive stress versus crack opening displacement, $\sigma - w$, curve, i.e., $G_F = \int_0^{w_c} \sigma(w) dw$. Different $\sigma - w$ relations were introduced since,

such as linear, bilinear, and exponential decay. The relation varies with materials and has to be determined experimentally.

2. Size-effect model – For members that follow the strength criterion or yield criterion, the log-log plot of peak loads versus member sizes represents a horizontal line, i.e., there is no size effect. The linear elastic fracture mechanics yields the strongest possible size effect which corresponds to a straight inclined line of slope $-1/2$. For materials that exhibit distributed cracking (fracture process zone), the size effect plot represents a smooth transition from the horizontal line for the strength criterion to the inclined straight line for the linear elastic fracture mechanics [Bazant and Pfeiffer, 1987]. Since test results vary with member sizes due to the size effect, it is difficult to find material fracture parameters. Thus, the size-effect model extrapolates the experimental data to those for the infinite specimen and based on that to define two fracture parameters G_f and c_f , where G_f and c_f are the critical energy release rate and the critical effective crack extension for the infinite specimen, respectively [Bazant and Kazemi, 1990]. To determine these two fracture parameters, at least three different sizes of geometrically similar specimens have to be tested. Bazant and Lin [1988] analyzed fracture tests on specimens of various sizes by finite elements using a nonlocal smeared cracking model. Their numerical results agreed well with the size effect law.

3. Two-parameter fracture model – As the load applied to a specimen increase from zero to its maximum value, the crack length changes from a_0 to a_c , where a_0 is the initial crack length and a_c is the critical effective crack length at the peak load. LEFM can then be used to present a criterion for failure: $\sigma_{Nc} = \frac{K_{Ic}}{F\sqrt{\pi a_c}}$, where

σ_{Nc} is the nominal failure stress (the peak value of the nominal stress), K_{Ic} is the critical stress intensity factor for the specimen, and F is a geometry- dependent function. It has been shown that the value of K_{Ic} remains constant for specimens of the same material. However, the value of a_c depends on specimen size and geometry, and cannot be used as a structure-independent material parameter. Jenq and Shah [1985] found that when structures of the same material fail, values of the elastic part of the crack tip opening displacement are approximately constant, regardless of the size of the structure. Based on their observation, they proposed the two-parameter fracture model, and claimed that the criterion for failure of a structure can be expressed as simultaneous satisfaction of $K_I(\sigma_{Nc}, a_c) = K_{Ic}^s$ and $CTOD(\sigma_{Nc}, a_c) = CTOD_c$, where $CTOD$ represents the elastic crack tip opening displacement at the location of initial crack tip. Functions for K_I and $CTOD$ on the left side of the above two equations in terms of applied load and crack length are based on LEFM and can be found in LEFM manuals for various structure geometries. Both K_{Ic}^s and $CTOD_c$ are regarded as material fracture parameters that are independent of structural size and geometry. The value of critical crack length a_c is determined by equating the unloading compliance of applied load versus crack mouth opening displacement at the peak load to the initial compliance, so its value is somewhere between the actual crack length and the length of the actual crack as well as the fracture process zone.

Implementation of New Fracture Models into FEINT and Critical Experiments

To improve the fracture behavior predictions of solid rocket motor firing, the current source code subroutines in FEINT related to fracture modeling should first be carefully reviewed if bugs exist. Next, some restrictions in the current implementation of FEINT should be removed. These restrictions include: the J-integral method assumes that steady state conditions exist around the crack tip, and that element stresses are in global coordinates; shell, beam, and truss elements are not allowed; internodal constraints are ignored and should not be used inside the outermost contour of elements. Furthermore, the actual material model should be used to determine the variation of strain energy density for computing J-integral values.

To incorporate Schapery's viscoelastic integral J_V into FEINT, the current procedure in finding J-integral values can still be used. However, the program has to be modified to include the option of computing J-integrals based on pseudoelastic displacement Δ^e . To determine the fracture toughness of propellant using J_V , stress relaxation tests have to be carried out on propellant specimens to find the effective Young's modulus, the time-averaged Young's modulus, and finally the load versus pseudoelastic displacement curve.

To incorporate Hillerborg's fictitious crack model, we can employ special spring elements along a crack to connect the two crack surfaces. The tension versus displacement curve of the spring element possesses an infinite initial stiffness following by a softening curve, which decreases from the point corresponding to the material's tensile strength. The softening curve is related to the cohesive stress distribution in the fracture process zone. When the stress in the spring is less than the tensile strength of the material, the stiffness is infinite, i.e., the crack is closed. If the stress goes beyond the tensile strength, it will then decrease with the increase of the crack opening displacement. The cohesive stress distribution of a propellant has to be determined experimentally.

To find the fracture parameters defined in Bazant's size-effect model, at least three different sizes of geometry similar specimens have to be tested. The recorded peak loads are then used to determine the fracture parameters by a curve fitting of the size-effect law formula. For Jenq and Shah's two-parameter fracture model, a three-point bending on a notched-beam is tested to record the applied load and the crack mouth opening displacement. The applied load is then intentionally unloaded when the load passes the peak value and is within the range of 95 percent of the peak load. The initial compliance, the unloading compliance, and the peak load are the quantities needed to calculate fracture parameters K_{Ic}^s and CTOD_c.

Discussions and Conclusions

The FEINT program has integrated many powerful modules into one unique system and greatly facilitates the analysis of solid rocket motors. It allows multiple coupled analyses, including burnback, internal ballistics, crack combustion, structural analysis, crack propagation, etc. Furthermore, it is a fully interactive, central-data-base-oriented code. All these features make it capable of performing a comprehensive structural-ballistic analysis of a

solid rocket motor firing, including an analysis of the growth of defects and their effects on motor operation. Probably, FEINT is the only software system available that possesses all these capabilities. The next step is to fine tune each individual component in the system so that quantitative numerical predictions can further be improved.

Several problems critical to a successful prediction of solid rocket motor performance should be addressed. Having adequate constitutive relations for the materials that form the motor (propellant, insulator, adhesive liner, and case) are essential. A linear elastic model is the simplest one. More complicated models can include plastic effects, viscoelasticity, and viscoplasticity. Temperature-dependent material properties and rate effect should also be included. Obviously, creep and sometimes stress relaxation become important considerations. Cracks, voids, and debonds (adhesive cracks) compound the difficulty of the modeling. Their existence alters the ballistics performance by exposing a larger burning surface area. Proper fracture criteria to determine when a crack will propagate and, if so, which direction and how far the crack will go have to be sought. The modeling of debond or adhesive crack propagation is even harder because it involves two different materials at a crack tip.

Schapery's J_V integral was particularly developed for viscoelastic materials. It is suggested to incorporate this integral into FEINT as a fracture toughness to predict if a crack will propagate. As to the influence of the fracture process zone, Hillerborg's fictitious crack model, Bazant's size-effect model, and Jenq and Shah's two-parameter fracture model can be considered. In the Thiokol's subscale motor test program, the crack propagation speeds were generally quite overpredicted. Liu and Ravi-Chandar [1996] showed that a power law relationship exists between mode I stress intensity factor and the crack growth speed. It was also observed in the Thiokol's test program that some motor failures were caused by splitting of the insulator, rather than the crack propagation in the propellant, at the aft flap tip, allowing the flame to burn through the aft dome. It is thus suggested that a future numerical analysis should take into account the effect of the sharp tip caused by making a flap in the insulator, or this sharp tip should be prevented during the motor manufacture.

It is worth mentioning that the FEINT program is not yet automated and not so user-friendly. It requires an expert analyst who is familiar with the knowledge of finite element modeling, structural analysis, ballistics, and fracture mechanics. The user has the flexibility to solve a variety of complex problems by having a complete control of the analysis sequences, including load step sizes, mesh refinement, and decisions on how to proceed based on the results returned from each individual analysis module. However, the user is also responsible for assuring that the approach and the procedure are correct and the results are reasonable. In addition, a three-dimensional analysis is time consuming, especially if a crack is propagating in a region of complex geometry, as in the analysis of the natural debond. Building a three-dimensional crack mesh is slow and difficult. This problem is intensified by the fact that analysis of a propagating crack requires the crack tip mesh to be refined each time the crack advances. More powerful three-dimensional crack meshing tools are needed.

Solid rocket motor firing is a complex process; its numerical simulation requires knowledge in various disciplines. The present state of the art in the problems of material constitutive relationship and crack and debond propagation modeling is still incomplete. A team work effort involving separate groups attacking specific

problems respectively is necessary. It is my believe that a precise prediction of solid rocket motor firing performance for structural ballistic risk assessment is achievable.

Acknowledgment

Thanks to Research & Development Laboratories, AFOSR Program Office and Phillips Laboratory, Edwards Air Force Base for their financial supports and providing me with the opportunity of conducting the summer research. Special thanks are due to Durwood Thrasher for his consistent hospitality and full supports throughout my summer research visit. Thanks are extended to Russell Leighton, Pat Empleo, and Robert Nichols for their friendship and technical assistance in the hot summer.

References

1. P. Anderson and L. P. Abrahamson, "Methods for Energy Release Rate Calculation, CPIA Report 41A, V.1, pp. 193-199, 1984.
2. T. L. Anderson, *Fracture Mechanics: Fundamentals and Applications*, 1st Ed., CRC Press, Boca Raton, FL, pp. 373-513, 1991.
3. Z. P. Bazant and M. T. Kazemi, "Determination of Fracture Energy, Process Zone Length, and Brittleness Number from Size Effect with Application to Rock and Concrete," *International Journal of Fracture*, V.44, No.2, pp. 111-131, 1990.
4. Z. P. Bazant and Feng-Bao Lin, "Nonlocal Smeared Cracking Model for Concrete Fracture," *ASCE, Journal of Structural Engineering*, V.114, No.11, pp. 2493-2510, 1988.
5. Z. P. Bazant and P. A. Pfeiffer, "Determination of Fracture Energy from Size Effect and Brittleness Number," *ACI Mat. J.*, V.84, pp. 463-480, 1987.
6. C. D. Bencher, R. H. Dauskardt, and R. O. Ritchie, "Microstructural Damage and Fracture Processes in a Composite Solid Rocket Propellant," *Journal of Spacecraft and Rockets*, V.32, No.2, pp. 328-334, 1995.
7. G. A. Collingwood, T. Sato, S. A. Geddie, A. M. Ratte', and E. B. Becker, "TEXLESP - 2D User's Manual," Morton Thiokol, Inc., and University of Texas at Austin, MSFC-RPT-1564, 1988.
8. T. C. Derbidge, "One-dimensional Compressible Flow in a Variable Area Duct With Mass Addition," Master's Degree Thesis, Department of Mechanical Engineering, University of Utah, Salt Lake City, Utah, 1971.
9. R. P. Graham, M. P. Iverson, D. P. Clark, D. A. Isaac, and J. B. Endicott, "Structural Ballistic Risk Assessment Methodology - Final Report for Period June 1988 to May 1995," Thiokol, Publication No. 951026, prepared for Phillips Laboratory, Edwards Air Force Base, 1995.
10. A. Hillerborg, M. Modeer, and P. E. Peterson, "Analysis of Crack Formation and Crack Growth in Concrete by Means of Fracture Mechanics and Finite Elements," *Cement and Concrete Research*, V.6, No.6, pp. 773-782, 1976.
11. Y. S. Jenq and S. P. Shah, "Two Parameter Fracture Model for Concrete," *Journal of Engineering Mechanics*, V.111, No.10, pp. 1227-1241, 1985.
12. C. T. Liu and K. Ravi-Chandar, "Local Fracture and Crack Growth Behavior in a Particulate Composite Material," *Journal of Reinforced Plastics and Composites*, V.15, pp. 196-207, 1996.
13. Y.-C. Lu, V. Yang, and K. K. Kuo, "User's Manual of Crack/Debond Combustion Anomaly (CDCA) Code," Pennsylvania State University, 1992.
14. D. M. Parks, "A Stiffness-Derivative Finite Element Technique for Determination of Crack Tip Stress Intensity Factors," *International Journal of Fracture*, V.10, pp. 487-501, 1974.
15. D. M. Parks, "The Virtual Crack Extension Method for Non-Linear Material Behavior," *Computer Methods in Applied Mechanics and Engineering*, V.12, pp. 353-364, 1977.

16. J. R. Rice, "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks," *Journal of Applied Mechanics*, V.35, pp. 379-386, 1968.
17. R. A. Schapery, "Correspondence Principles and a Generalized J-integral for Large Deformation and Fracture Analysis of Viscoelastic Media," *International Journal of Fracture*, V.25, pp. 195-223, 1984.
18. R. A. Schapery, "On Some Path-Independent Integrals and Their Use in Fracture of Nonlinear Viscoelastic Media," *International Journal of Fracture*, V.42, pp. 189-207, 1990.

Table 1: Subscale Motor Tests

Test No.	Test Date	Description	Nozzle Throat, ϕ	FEM Model
1	5/17/93	Original baseline, no flaps, no flaws	2.0 in.	Axisymmetric
2	7/13/94	Redesign baseline, with flaps, no flaws	2.3 in.	Axisymmetric
3	7/14/94	Full-length bore crack, 0.3 in. deep	2.3 in.	Gen. Plane Strain
4	7/19/94	Full-length bore crack, 0.35 in. deep	2.0 in.	Gen. Plane Strain
5	3/13/95	Buried debond, axisym., 0.75 in. wide	2.3 in.	Axisymmetric
6	3/14/95	Buried debond, axisym., 0.75 in. wide	2.3 in.	Axisymmetric
7	3/15/95	Voids, 0.72 x 0.47 x 0.30 in. and smaller	2.3 in.	3-D
8	3/16/95	Edge debond, axisymmetric, 1.0 in. deep	2.3 in.	Axisymmetric
9	3/17/95	Edge debond, axisymmetric, 1.0 in. deep	2.3 in.	Axisymmetric
10	3/17/95	Half-circle plane bore crack, $r = 0.4$ in.	2.1 in.	3-D
11	3/20/95	Half-circle plane bore crack, $r = 0.6$ in.	2.1 in.	3-D
12	3/22/95	Half-circle plane bore crack, $r = 1.7$ in.	2.1 in.	3-D
13	3/27/95	0.3-in. axisym. edge and natural debonds	2.3 in.	3-D

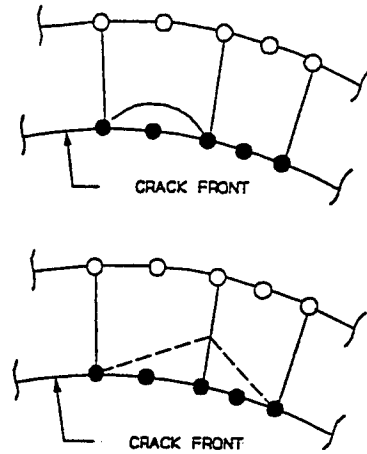


Fig. 1 Crack Tip Perturbation for Computing J-integral.

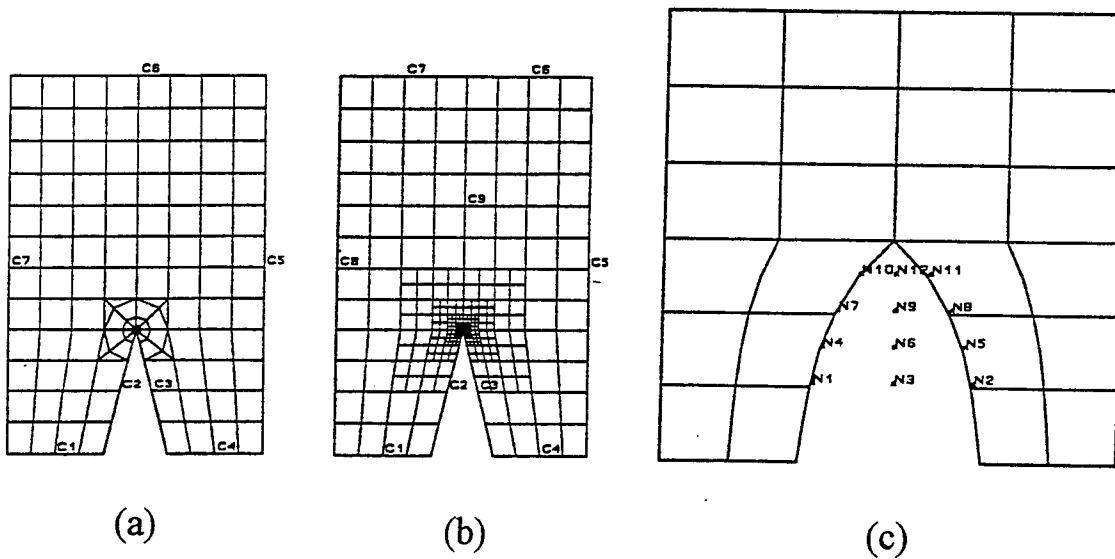


Fig. 2 Three Crack Tip Mesh Types Available in FEINT: (a) Focused Mesh, (b) Rectangular Mesh, and (c) Zipper Mesh.

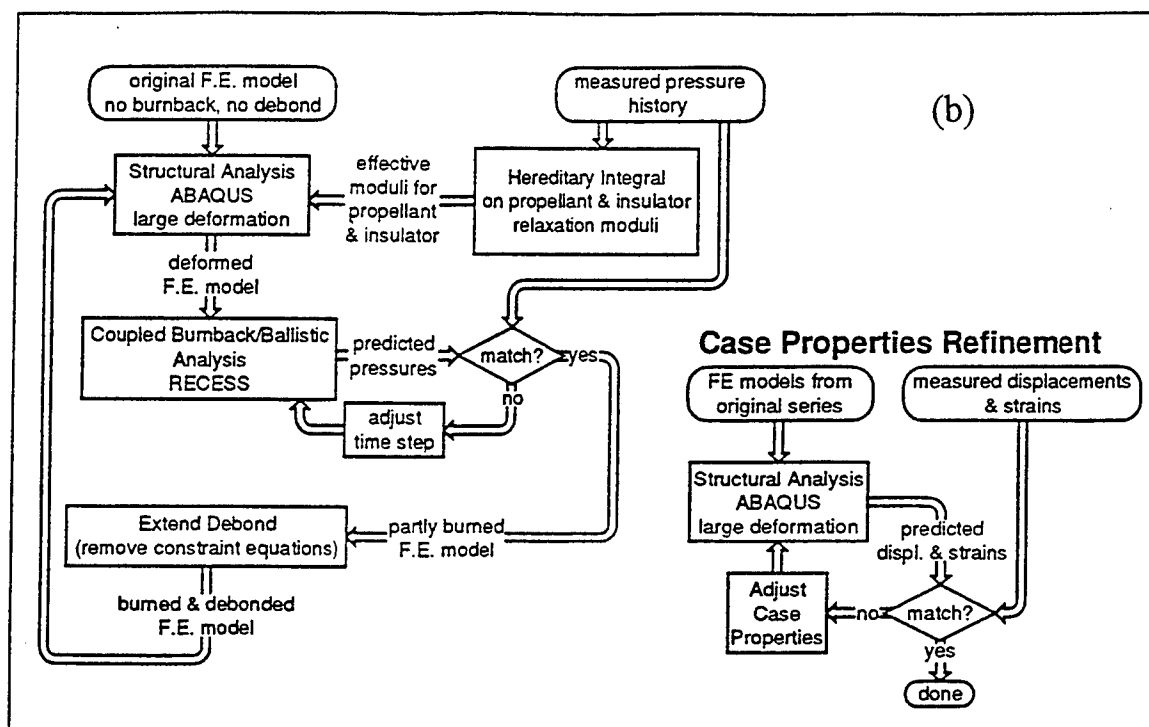
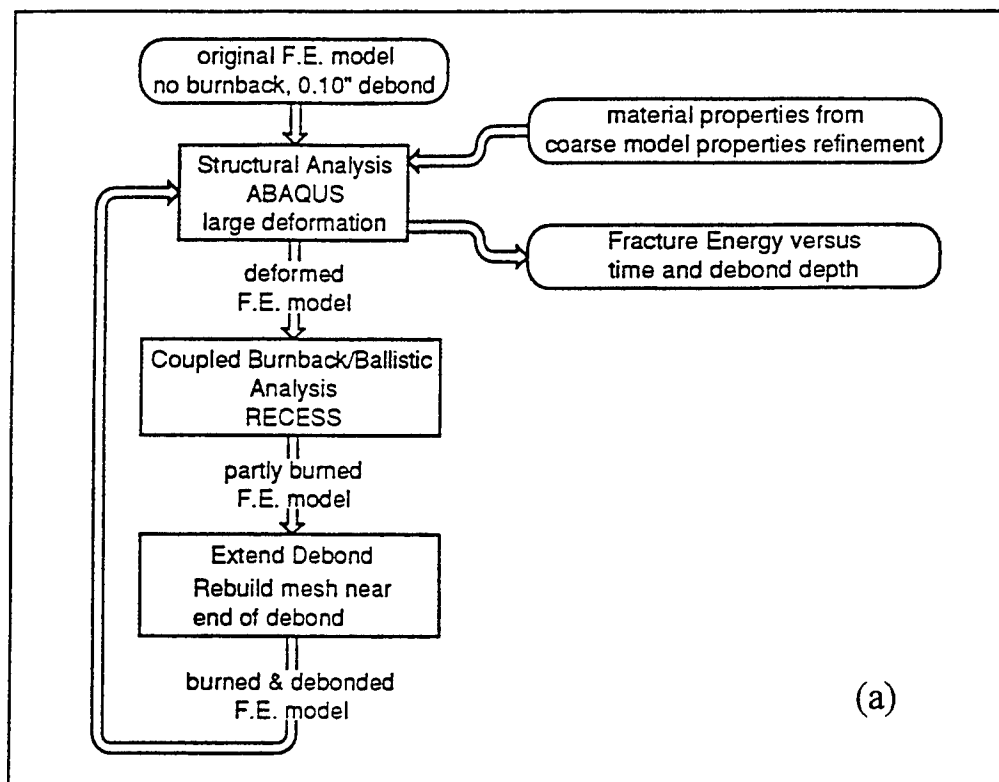


Fig. 3 Numerical Analysis Flow for Baseline Motor Analysis: (a) Coarse Model Analysis Flow, (b) Fine Model Analysis Flow.

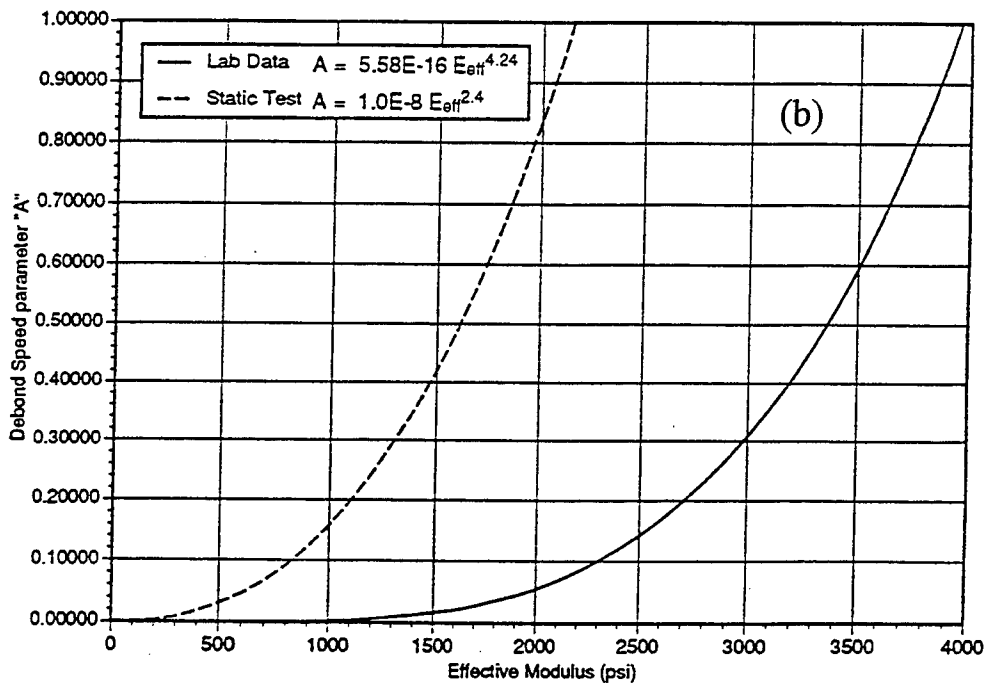
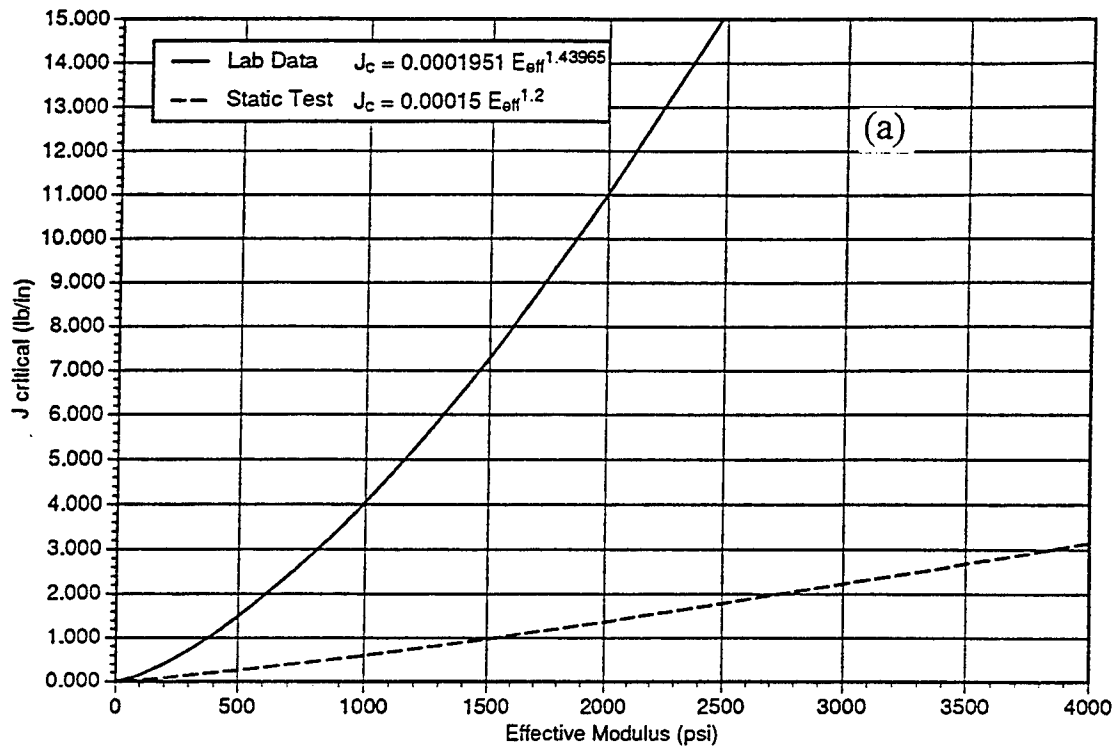


Fig. 4 Fracture Parameters Obtained from Motor Firing Tests and Laboratory Specimen Tests: (a) Critical J-integral Value versus Effective Modulus, (b) Debond Speed Parameter A versus Effective Modulus.

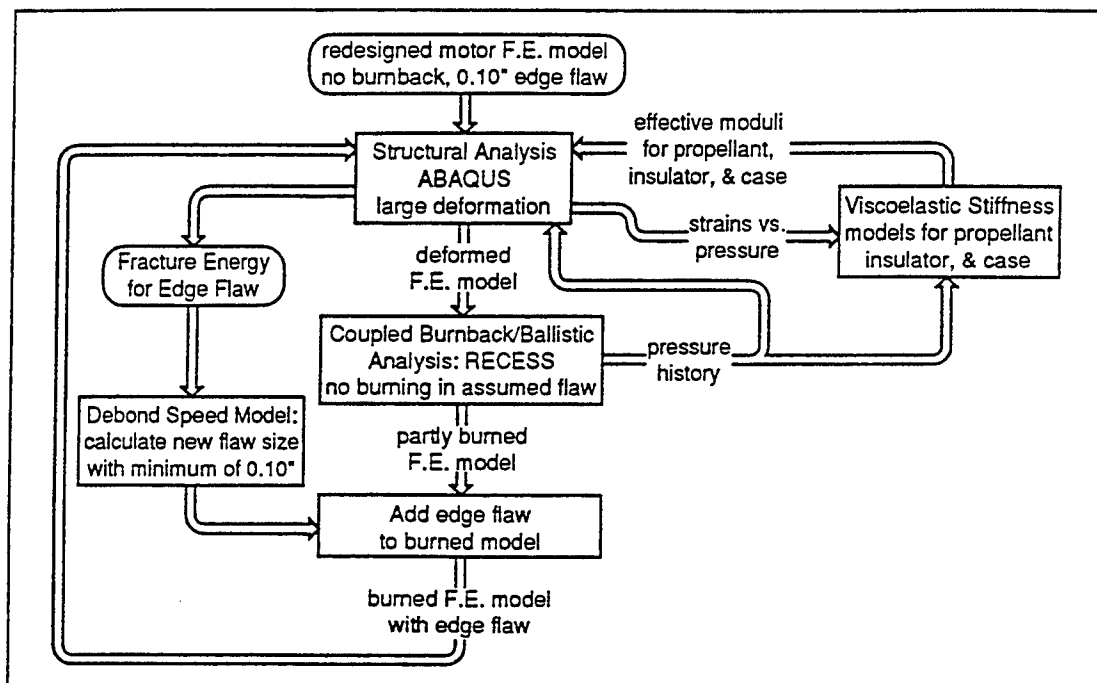


Fig. 5 Numerical Analysis Flow for Redesigned Motor Analysis.

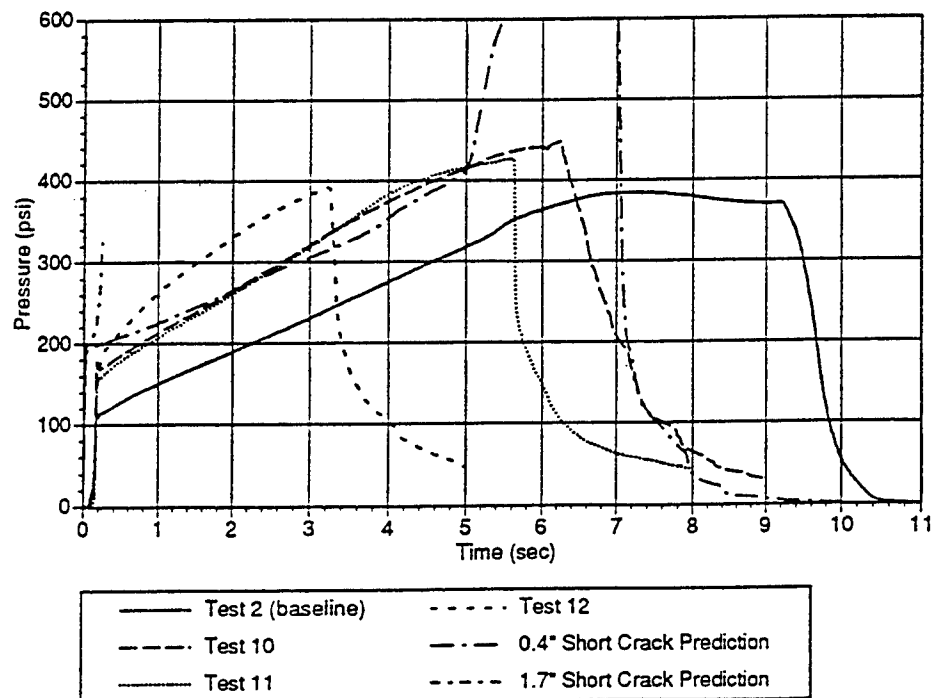


Fig. 7 Numerical Prediction and Test Measurement for Subscale Motor Test 10, 11, and 12.

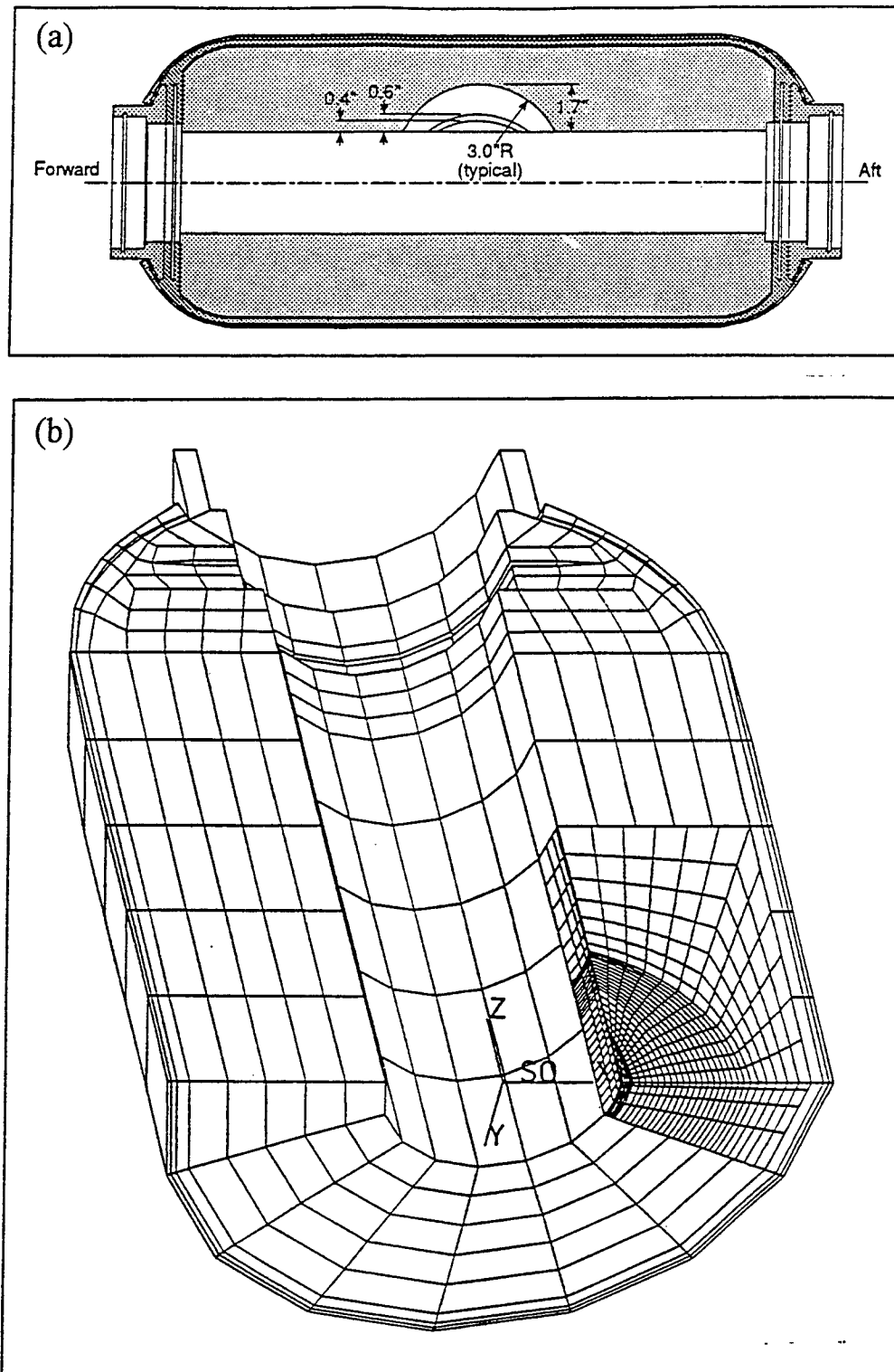


Fig. 6 Finite Element Mesh with Half-Circle Plane Bore Crack: (a) Bore Crack Configuration, (b) 3-D Finite Element Mesh.

THEORY, MODLING AND ANALYSIS OF AMTEC

M. A. K. Lodhi
Professor
Department of Physics and Engineering Physics

Texas Tech University
Lubbock, Texas 79409

Final Report for
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

July 1996

THEORY, MODELING AND ANALYSIS OF AMTEC

M. A. K. Lodhi

Professor

Department of Physics and Engineering Physics
Texas Tech University

Abstract

Alkali-Metal Thermal-to-Electrical Converter (AMTEC) is a high temperature regenerative concentration cell for elemental sodium which converts thermal energy directly into electrical energy. The efficient operation of AMTEC cell involves several challenging heat and mass transfer processes which require some through investigation. An actual AMTEC cell consists of several sodium vapor tubes. In order to understand its working principle and develop a good working model we first started our analysis with a single-tube device. For that a single-tube AMTEC cell is divided into four flow nodes. The pressure drop is calculated for each node in order to calculate the total pressure drop through the tube. The flow equations are derived in continuum, free molecular motion and transition regimes. As expected the molecular flow regime is more challenging. We studied this regime with three different approaches. None of them agrees with the observed data in the high current region. A fourth option, called Dusty Gas Model (DGM) is introduced . This approach is expected to remove the discrepancy between the observed data and the results predicted by this model. If so, this model will be applied to multi-tube cell.

THEORY, MODLING AND ANALYSIS OF AMTEC

M. A. K. Lodhi

1. Introduction

The conversion of thermal and/or mechanical energy to electrical energy has been historically the mainstay for power systems. With the advent of Faraday's law of electromagnetic induction and the steam engine, dynamic conversion systems evolved rapidly and were perfected with engineering details mostly for terrestrial and aircraft devices. As a potential alternative to dynamic conversion systems, static thermal to electrical conversion systems have been investigated, particularly for space power systems. With the discovery of the Seebeck effect, the thermoelectromotive force, which occurs in materials under the effect of a temperature gradient, has been the driving force for static thermoelectric generators. A large number of static devices have been evolved and studied. The goal for static conversion systems has been to identify static converters with efficiencies that are competitive with dynamic systems. To this end currently, NASA Lewis Research Center (LeRC) has an interest in developing thermal energy storage (TES) for the solar dynamic ground test demonstration (SDGTD) program. Phillips Laboratory has an interest in developing an electrical power system for the payload in Low Earth Orbit (LEO) for a duration of 5 years. For conversion technology, Jet Propulsion Laboratory (JPL) is currently interested in alkali metal thermal to electrical conversion (AMTEC) with respect to the Pluto Express Mission. In all these endeavors the effort is aimed at developing advanced power conversion and thermal management and control subsystems for space applications. As part of this effort, Alkali-Metal Thermal-to-Electric Converter (AMTEC) device is being developed and tested at USAF Phillips Lab with the goal of providing feedback or improved efficiency, durability, and long life time. To reach this goal, thorough and accurate mathematical modeling and analyses as well as experimental testing of AMTEC device is required. Our

efforts therefore, on theoretical side have been to develop an accurate mathematical model of AMTEC at Phillips Lab this summer.

2. Background of AMTEC

AMTEC is a relatively new type device, based on the principle of sodium concentration cell, conceived in late sixties [1,2]. AMTEC devices are a leading power conversion candidate for a large number of applications, particularly low mass, low cost, long life, and high-efficiency power systems for satellites and exploratory space probes. AMTEC systems are very attractive for these applications, because they are static conversion devices which can provide efficiencies close to the theoretical Carnot efficiency. Efficiencies as high as 19% have been measured in laboratory experiments. Optimized AMTEC devices of a mature design can provide theoretical conversion efficiencies between 20 and 40% when operating at a moderate hot-side temperature between 900 K and 1300 K and a cold-side temperature between 400 K - 800 K. Other advantages of AMTEC devices include low maintenance, high durability, high specific power, modular construction, and the ability to operate in conjunction with a high-temperature combustion, a nuclear or a solar heat source. Therefore, AMTEC devices are attractive not only for space, but also for many terrestrial and aerospace applications.

2.1 Operation Principles of AMTEC Cell

To understand the working of AMTEC, its operating cycle is illustrated in Fig 1 [3]. A closed vessel is divided into a high temperature (T_2) and high pressure (P_2) region in contact with a heat source and a low temperature (T_1), low pressure (P_1) region in contact with a heat sink. AMTEC is a high-temperature regenerative concentration cell for elemental sodium, which directly converts heat to electrical energy. The critical material in the operation of the AMTEC is the beta"-alumina solid electrolyte (BASE), a sodium-ion conductor whose ionic conductivity is much larger than its electronic conductivity. In an AMTEC cell, the BASE separates the hot (high-pressure) region filled with a small quantity of liquid sodium (typically <10%) in contact with the heat source, from the cold (low-pressure) region occupied

Also inner side for vapor-vapor cell.

with sodium vapor. A porous electrode covers the low-pressure (outer) side of the BASE. Electrical leads in contact with the porous electrode and the high-temperature liquid sodium exit through the wall of the device and are connected to an external load. The pressure differential across the BASE forces ionization of sodium atoms on the hot-side. The ions diffuse through the BASE toward the low-pressure side in response to the pressure differential gradient of free Gibbs energy while the electrons circulate through the external load, producing electrical work [4]. Electrons and sodium ions recombine at the interface between the BASE and the porous electrode. The resulting sodium atoms absorb their heat of vaporization, move through the electrode and the vapor space, then release their heat of vaporization at the low-temperature condenser surface. A wick structure or an electromagnetic pump return the liquid condensate to the high temperature side of the BASE to complete its circulation. More details on AMTEC are in scarce literature [1-6]. AMTEC possesses conversions efficiencies much higher than other direct thermoelectric devices. It was reported at 19% in 1981 [7], with a power density of 1 W/cm^2 [6]. Its hot region temperature is in the range of 900-1300 K [8]. This makes it compatible with the general purpose heat source (GPHS) [9] or the projected SP-100 nuclear space power reactor [10]. AMTEC is limited to 1350K because of sodium interactions with the BASE. The lower temperature region is limited to 500 K by the need to maintain the sodium in the liquid state. As the condenser temperature increased above 700 K the efficiency decreases [11]. Besides, AMTEC has high efficiency near Carnot [12], low maintenance and high durability, modular construction and the ability to use high temperature combustion, nuclear, or solar heat source [13].

*? minimize
vapor back
pressure*

3. Heat and Mass Transfer Processes to be Investigated

The operation of an AMTEC cell involves several challenging heat and mass transfer processes requiring thorough investigations, namely: (a) the diffusion of sodium ions through the BASE tube; (b) the evaporation of sodium at the BASE-porous electrode interface; (c) the sodium vapor flow through the porous electrode and the vapor space in the free-molecular regime; and (d) the recirculation of liquid sodium through the capillary tube. Any of these processes can limit the operation of the AMTEC cell. Because sodium is frozen at room temperature, the startup of the AMTEC cell from a frozen state is also a

concern. Of particular interest is the status and the effect of the initial distribution of the sodium working fluid in the cell.

The performance and the conversion efficiency of an AMTEC cell depend on the internal and external parasitic heat and electrical losses. The internal heat losses are made up of 1) conduction through the output current leads and 2) supporting structure of the BASE tube, and of internal radiation heat losses from the 3) porous electrode surface to the condenser surface (through the vapor space). The latter depends on the a) geometry, b) axial temperature distribution in the condenser shell and the c) emissivity of both surfaces. 4) Of particular importance is the fraction of condenser surface covered with the sodium condensate. If the porous electrode and condenser surfaces exchange radiative energy at rates comparable to those of ideal bodies, the maximum efficiency of the cell would be less than 10%. Fortunately, the reflectivity of the sodium liquid film is greater than 98% in the infrared. Thus, if the surface of the film is smooth, the radiative losses could be as low as 2% of black body emission, even if the porous electrode is an ideal emitter. To take advantage of the high specular reflectivity of the sodium condensate, the shape of the condenser surface must be designed to reflect most of the radiation emanating from the porous electrode back to it.

The external heat losses are made up of conduction through the 5) support structure of the condenser (or the outer shell of the cell), and 6) radiation from the cell outer shell to the shield surrounding the cell. 7) Heat leakage through the radiation shield may also be important. (In the case of a system comprised of several AMTEC cells, radiation exchange between the shells of the different cell must be considered.) Due to the high operating temperature of the AMTEC radiation is a major heat transfer mechanism. 8) A major challenge involved in calculating radiation heat transfer is determining the **geometric view factor**. 9) Except for the sodium vapor there is no convection heat loss. The sodium vapor heat transfer through convection only affects the internal degree of freedom of AMTEC. Therefore, the AMTEC thermal model will mainly consider conduction and radiation heat losses. Because AMTEC cells for space applications would operate in a vacuum, convection heat transfer is non-existent. For this reason, the AMTEC cells are tested in vacuum chambers.

4. Research Tasks

Phillips Laboratory is planning the design, testing and analysis of several AMTEC cells, in collaboration with Advanced Modular Power Systems, Inc. (AMPS), located in Ann Arbor, Michigan. The AMTEC cells constructed by AMPS will be tested in a vacuum chamber and results will be analyzed. Two different types of cells are being considered for testing, single-cell units and multi-cell units for the Pluto Express Mission. For the purpose of this research period we were primarily concerned with the single cell unit. The model thus developed will be applied to multi-tube unit.

We formulated a mathematical model of AMTEC which takes into consideration all those aspects, discussed in section 3, thus leading to a prototype designing, fabricating, testing and further improving of the cell. The three flow regimes, just mentioned above, are arbitrarily determined by Knudsen number, Kn , as follows [14]:

(a)	Continuum (viscous)	$Kn \leq 0.01$
(b)	Slip (transition)	$0.01 < Kn < 1$
(c)	Free Molecular (molecular)	$Kn \geq 1.$

These situations are illustrated in Fig. 2.

This necessarily requires to divide the AMTEC cell, under investigation into appropriate parts as accurately as possible where those regimes dominate and derive equations necessary to calculate the vapor flow pressure drop including those regimes.

This model has four options for the effect of the molecular collisions with the wall particularly in the rarefied flow regime. In order to incorporate the effect of molecule- wall collisions more realistically, the so called Dusty Gas Model (DGM) is used as the fourth approach [15-18]. The DGM treats this effect as a diffusion problem involving the original gas molecules, in motion, and the giant wall molecules at rest and uniformly distributed. The problem is thus formally treated as the one component flow of a multicomponent gas mixture, as shown schematically in Fig. 3.

The vapor flow pressure drop in various parts of AMTEC representing the aforementioned regimes has been compared with the observed data taken from the device under investigation.

5. Conclusion

The literature offers several options for determining form loss coefficients. We tried the computer simulation model for three options. We ran the computer model results for two options. The third one is in progress. The first option is for laminar flow with Reynolds number, $Re < 10$. The other option considered is to assume that the heat shield acts like a thick edged orifice between two pipes of different diameters [19]. The pressure drops for these options are given in Figs. 4 and 5. The computer code is given in Appendix A. The DGM is still in progress. We expect that the DGM will remove the kink in the pressure drop at current (mass flow) = 4 A and the discrepancy with the observed data at high current.

References

1. R.C. Dahlbert et.al, "Review of Thermionic Technology 1983 to 1992," "A Critical Rev. of Space Nucl Power and Propulsion 1984 - 1992", Ed. M.S. El-Genk (AIP Press, NY 1994) pp 121-161
2. J.T. Kummer and Weber, U.S. Patent 3,458,356 (1968), Assigned to Ford Motor Co
3. G.A. Johnson (1994) "Study of the Rarefied Sodium Vapor Flow in the Pluto Fast Flyby AMTEC Cell," 11th Symposium on Space Nuclear Power and Propulsion , held in Albuquerque, NM, January 1994, AIP Conference Proceeding N^o. 301, M.S. El-Genk and M.D. Hoover, Eds., 2: 581-585
4. T. Cole, "Thermoelectric Energy Conversion with Solid Electrolyte", Science, 221,915 (1983)
5. N. Weber "A Thermoelectric Device Based on Beta-Alumina Solid Electrolyte", Energy Conversion 14, 1 (1974)
6. T.K. Hunt, N. Weber, and T. Cole, "Research on the Sodium Heat Engine", Proc 13th Intersoc Energy Conversion Engineering Conf., SAE, Warrandale, PA (1978) p201
7. T.K. Hunt, N. Weber, and T. Cole, "High Efficiency Thermoelectric Conversion with Beta" - Alumina Solid Electrolyte: The Sodium Heat Engine", Fast Ionic Transpost in Solids, ed J.B. Bates, G.C. Ferrington (N-Holland Co, 1981), P263
8. C.P. Bankston, T. Cole, S.K. Khanna and A.P. Thakoor, "Alkali Metal Thermodynamic Conversion (AMTEC) for Space Nuclear Power Systems", Space Nucl Power Sys. Ed. M.S. El-Genk and M.D. Hoover, (Orbit Co., Malabar, Fl, 1985) pp 398-402
9. A. Shock, "Design Evolution and Verification of the General Purpose Heat Source," Proc 15th Intersocieties Energy Conversion Engineering Conf Vol 2 (AIAA, NY 1980) p 1032
10. D. Birden and F.A. Angela, "Space Reactor-Past, Present and Future", Proc 18th Intersocieties Energy Conversion Engineering Conf, Vol (AICHE, N.Y., 1983) p 61

11. R. Ewell and J. Mondt "Static Conversion Systems", Space Nuclear Power Sys., Ed. M.S. El-Genk and M.D. Hoover (Orbit Book Co. Malabar, Fl., 1985) pp 385-391
12. C.B. Vinning, R.M. Williams, M.L. Underwood, M.A. Ryan, and J.W. Suiter, "Reversible Thermodynamic Cycle for AMTEC Power Conversion", Proc 27th Intersocieties Energy Conversion Engineering Conf., Society of Automobile Engineers, 1992, pp 3: 123-128
13. M.L. Underwood, et al, "Recent Advances in AMTEC Recirculating Test Cell Performance,"Conf 930103 (AIP, 1993) pp 885-890
14. S. Dushman and J. M. LAFFERTY, *Scientific Foundations of Vacuum Technique* (John Wiley & Sons, Inc., New York, 2nd Edition) 1962, p
15. R. B. Evans III, G. M. Watson and. A. Mason, " Gaseous Diffusion in Porous Media at Uniform Pressure," J. Chem. Phys., **35**, 2076-2083, (1962)
16. R. B. Evans III, G. M. Watson and E. A. Mason, " Gaseous Diffusion in Porous Media - II. Effect of Pressure Gradients, " J. Chem. Phys., **36**,1894-1902 (1962).
17. E. A. Mason, A. P. Malinauskas and R. B. Evans, J. Chem. Phys., **46** 3199-3216 (1967).
18. R. E. Cunningham and R. J. J. Williams, "Diffusion in Gases and Porus Media, " (Plenum Press, New York, 1980) p 131
19. J. F. Ivanenok III, " Modeling of Sodium Vapor Flow in Remote Condensing Alkali Metal Thermal to Electrical Converters, " M. S. Thesis, Univ. of Michigan (unpublished) 1994.

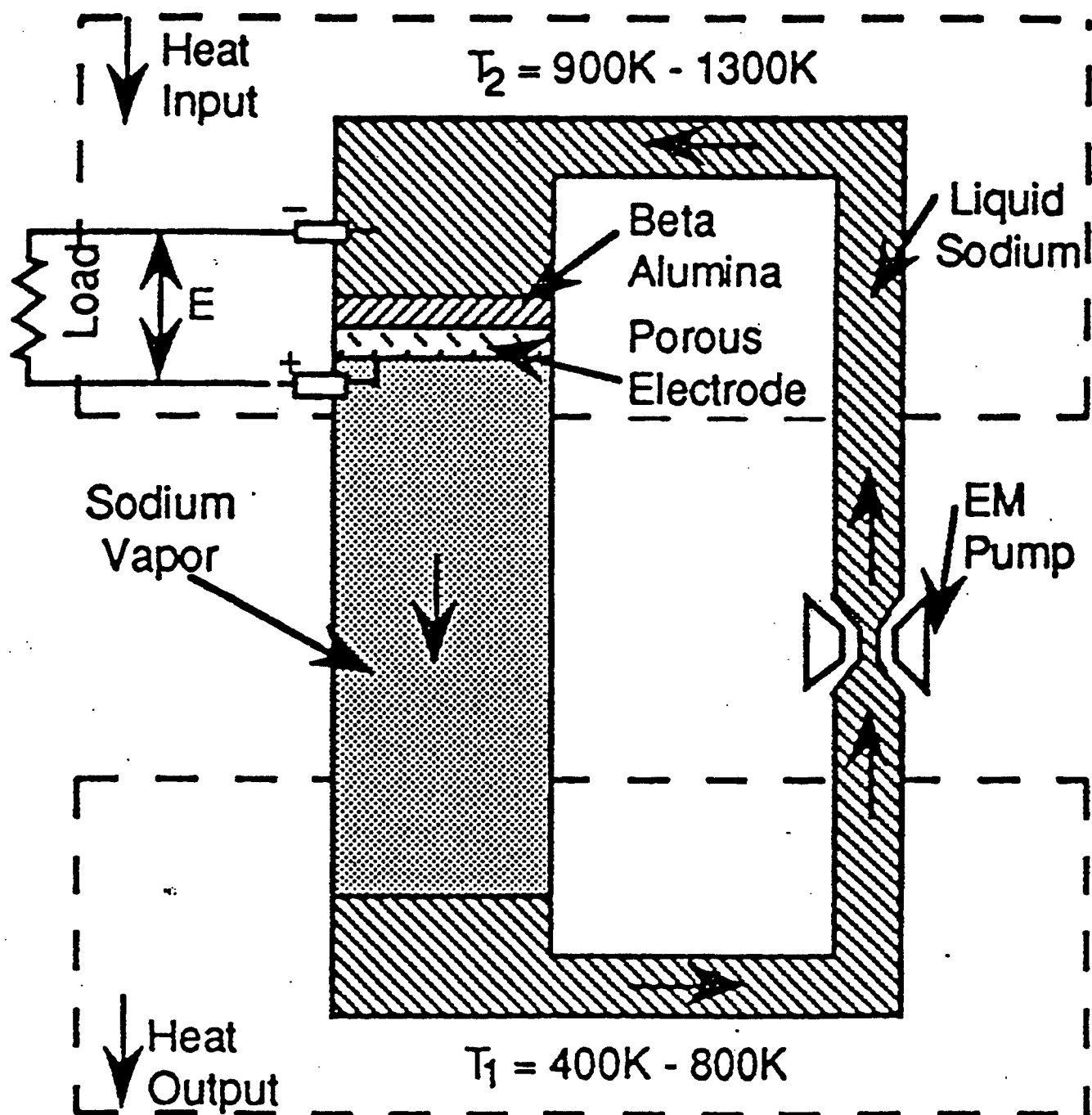
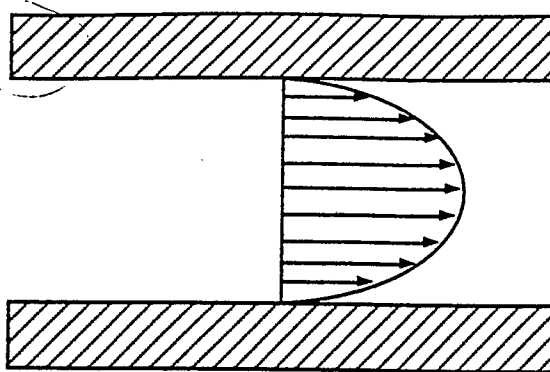
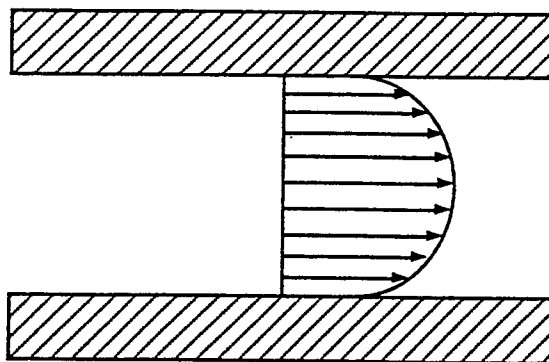


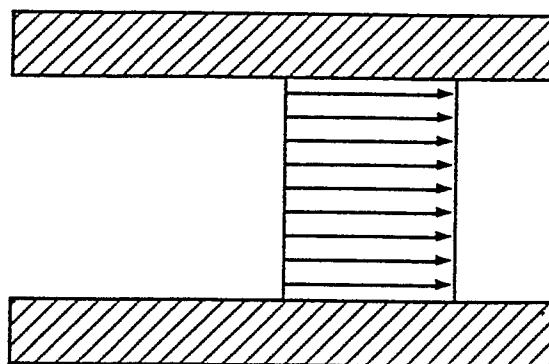
Fig. 1 Schematics of AMTEC



(a) viscous flow regime

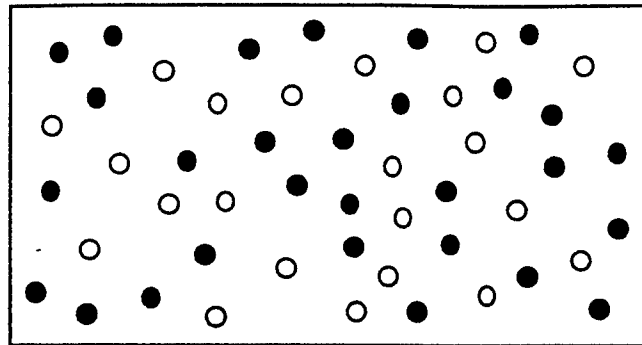


(b) transition flow regime



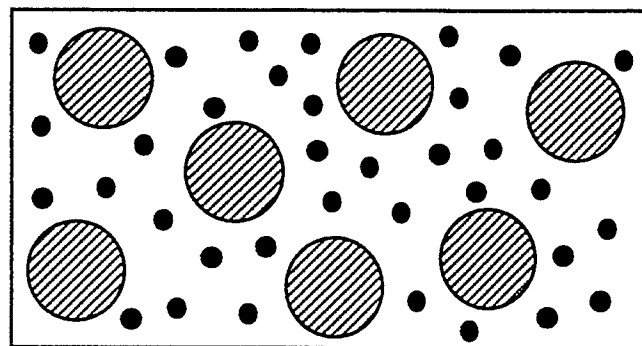
(c) free-molecule flow regime

FIGURE: 2 Illustration of Velocity Profiles in a Pipe.



- gas molecule (isotope 1)
- gas molecule (isotope 2)

(a) self-diffusion model

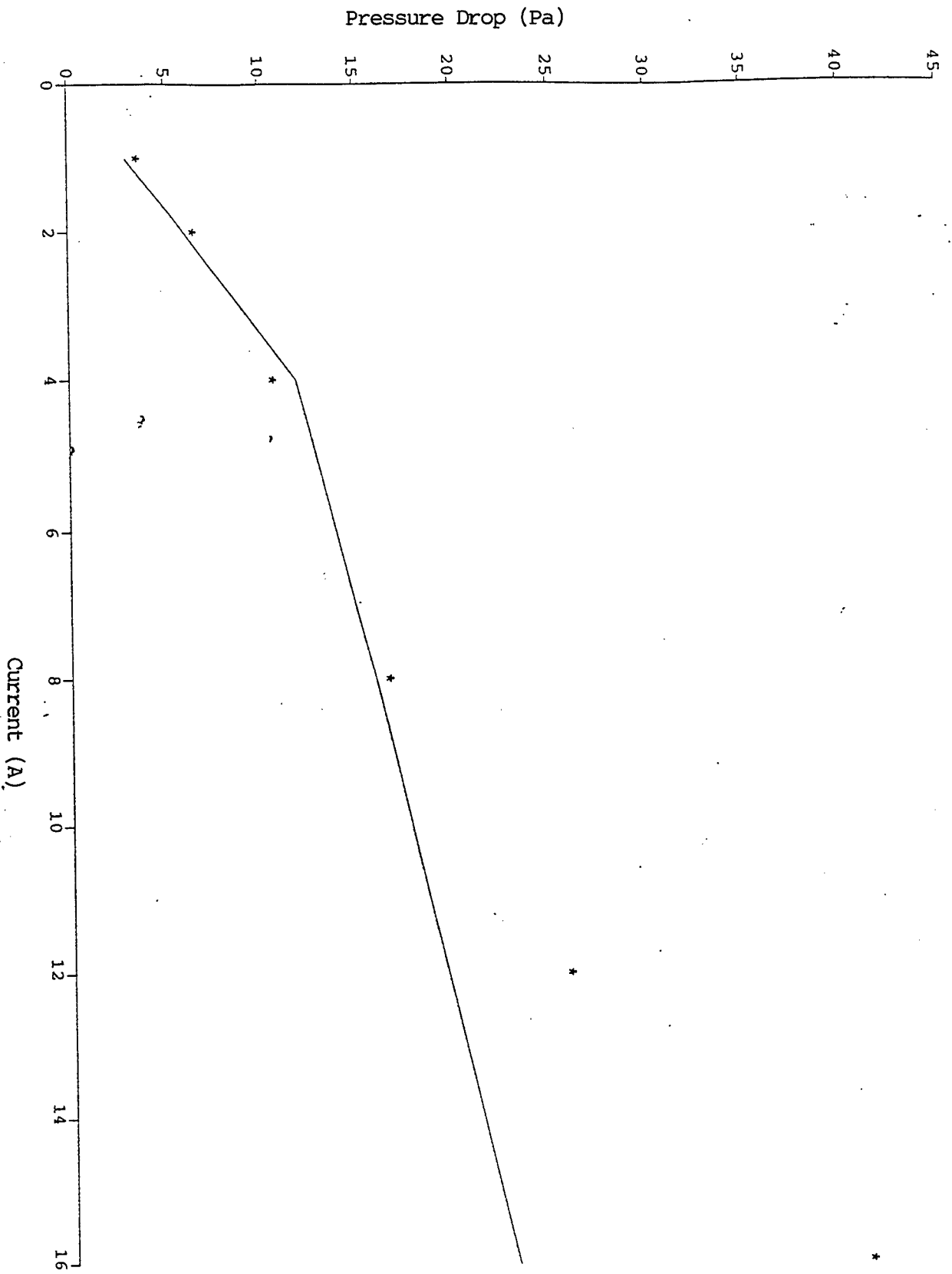


- gas molecule
- ◐ wall molecule (dust)

(b) Dusty Gas Model

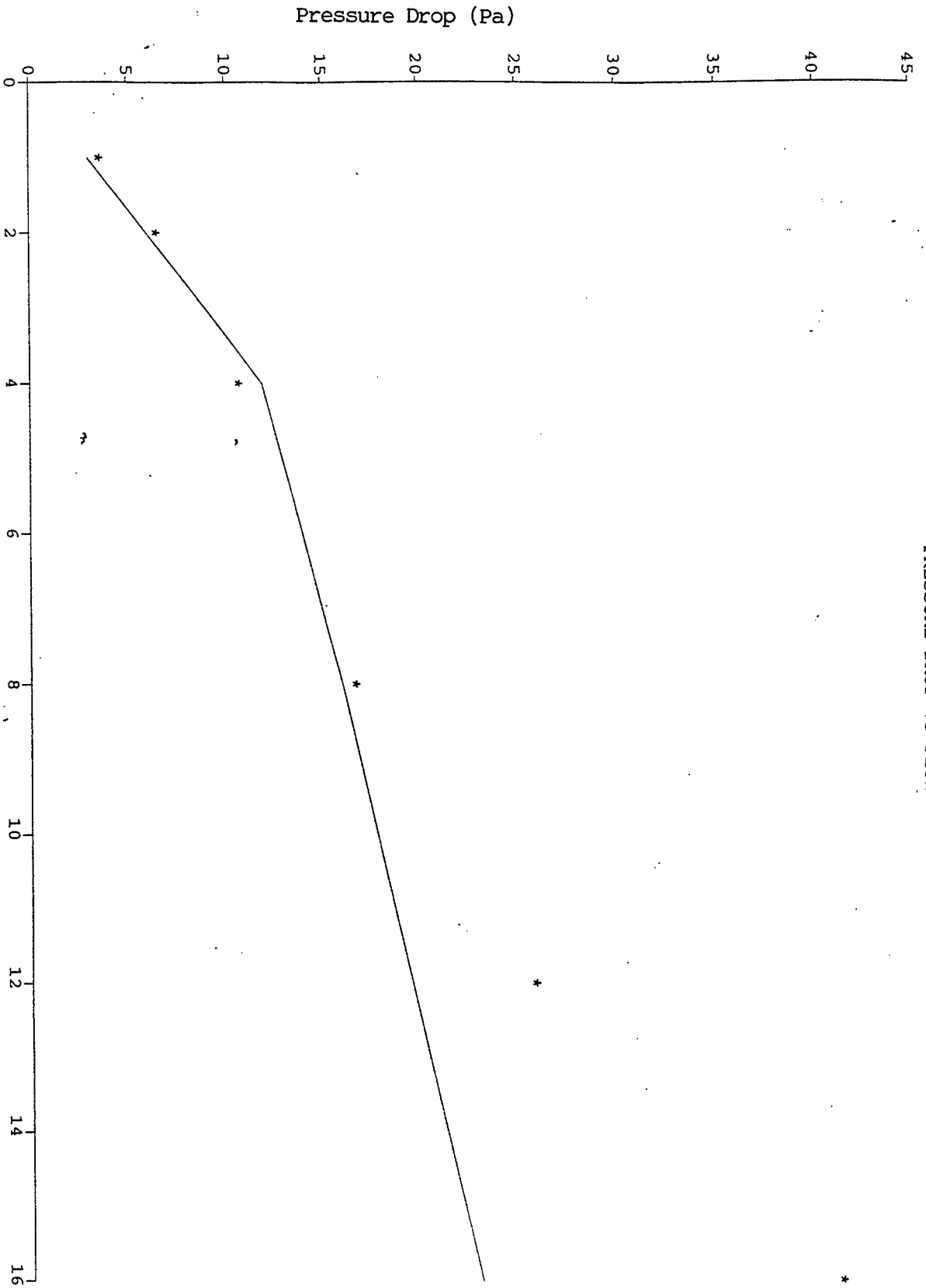
FIGURE: 3 Illustrations of Flow Models.

PRESSURE DROP VS FLOW



Current (A)
Fig. 4 Re<10 Option

PRESSURE DROP VS FLOW



Current (A)
Fig. 5 Thick Edge Option

;RC1000 Pressure Loss Model

* $\tilde{a}=\pi()$

;CELL DIMENSIONAL CALCULATIONS

* $L_b=LB1+Le1+LB2+Le2+LB3$

* $Condid = Dcond-2*tcond$

* $Condr = Condid/2$

* $Lcond = Ltc+Lc/2$

* $Cid = Dcel-2*tc$

* $Phs=\tilde{a}*(Dcond+Dhs)$

* $Dhhs=4*Ahsa/Phs$

* $Dhan = Cid-Dt$

* $Aan = \tilde{a}/4*(Cid^2-Dt^2)$

* $Ate = \tilde{a}*Dt*Le1$

* $Ahsa = \tilde{a}*Condid*Lhs$

* $Abhs = \tilde{a}/4*Cid^2-\tilde{a}/4*Dhsp^2$

* $Ac = \tilde{a}*Condid*Lc$

* $Acf = \tilde{a}/4*Condid^2$

* $I = j*Ate$

* $Voc2c=Voc2d+I*Rcor$

* $dT = Tb2-Tb$

* $Fé=(2-é)/é$

; SATURATION PRESSURE

* $\ln(Pb/101325) = 14.631-12818/Tb-0.5*\ln(Tb)$

* $\ln(Pb2/101325) = 14.631-12818/Tb2-0.5*\ln(Tb2)$

* $\ln(Pc/101325) = 14.631-12818/Tcond-0.5*\ln(Tcond)$

;ELECTRODE 2 MODEL

* $f=F/R/Tb$

* $f2 = F/R/Tb2$

* $Voc2c = \ln(Pb2/Pe2c)/f2$

* $dPcdd = Pe2c-Pc$

* $dPcm = Pe2-Pc$

* $dPerr = ((dPcdd-dPcm)/dPcdd)*100$

;FROM LOSS COEFFICIENTS FOR PRESSURE DROP MODELS

;FROM Frank M. White

S Rule

$$C \ K1 = (1-(Dhan^2/Cid^2))^2$$

$$C \ K2 = 0.42*(1-(Ahsa^2/Abhs^2))$$

$$C \ \text{if } Ahsa > Acf \text{ then } K3 = 0.42*(1-(Acf^2/Ahsa^2)) \text{ else } K3=0$$

$$C \ \text{if } Ahsa > Acf \text{ then } K4=0 \text{ else } K4 = (1-(Ahsa^2/Acf^2))^2$$

;From Loss From Idel'Chik

$$C \ K1 = 30/Recan$$

$$C \ K2 = 30/Rechs$$

$$C \ \text{if } Ahsa > Acf \text{ then } K3= 33/Recc \text{ else } K3=0$$

$$C \ \text{if } Ahsa > Acf \text{ then } K4=0 \text{ else } K4 = 30/Rechs$$

$$* \ Kt = K1+K2+K3+K4$$

$$* \ mdot = (j/F)*M*Ate$$

;Joe's Idea

$$C \ K1 = (30/Recan)*((1-(Dhan^2/Cid^2))^2)$$

$$C \ K2 = (30/Rechs)*0.42*(1-(Ahsa^2/Abhs^2))$$

$$C \ \text{if } Ahsa > Acf \text{ then } K3= (30/Recc)*0.42*(1-(Acf^2/Ahsa^2)) \text{ else } K3=0$$

$$C \ \text{if } Ahsa > Acf \text{ then } K4=0 \text{ else } K4 = (30/Rechs)*(1-(Ahsa^2/Acf^2))^2$$

;Prof. Schultz's Idea

$$* \ K1=0$$

$$* \ Lbar=0.5*(Dhs-Dcond)/Dhhs$$

$$* \ pf=0.25+0.535*Lbar^6/(0.05+Lbar^8)$$

$$* \ \text{if } Lbar \leq Lcri \text{ then } Tau = (2.4-Lbar)*10^{(-pf)} \text{ else } Tau = 0$$

$$* \ K2=(0.5*(1-Ahsa/Abhs)^{.75}+(1-Ahsa/Acf)^2+Tau*(1-Ahsa/Abhs)^{.375}*(1-Ahsa/Acf)+fcfh$$

$$* \ K3=0$$

$$* \ K4=0$$

;PRESSURE DROPS

$$* \ dPcc = (2*\tilde{a}*M*R*Tcond)^{.5}*j*(Ate/Ac)/F$$

$$* \ Pe2 = Pc + dPan + dPbhs + dPhs + dPcond + dPcc$$

$$* \ Rhoc=((Pc+dPcc)/(R*Tb))*M$$

;CONDENSER CONTINUUM FLOW PRESSURE DROP MODEL

$$* \ u = (1.6E-7*(Tb)-5.0E-6)/10$$

$$* \ Rhonavc=((Pc+dPcc+dPcnd)/(R*Tb))*M$$

$$* \ vcfc = mdot/(Acf*Rhonavc)$$

$$* \ Machc = vcfc/((9/7)*R*Tb/M)^{.5}$$

$$* \ Recc = (Rhonavc*vcfc*Condid)/u$$

S Rule

* $fcfc = 64/(Recc*(8*Knc*Fé+1))$
* $Alphac = (2+8*Knc*Fé+16*Knc^2*Fé^2)/((4*Knc*Fé+1)^2)$
* $PRc = (Pc+dPcc)/(Pc+dPcc+dPcnd)$
* $Alphac/n*\ln(PRc) = (n/(n+1))*((PRc^{(1/n+1)-1}*R*Tb)/(M*vcfc^2) + 0.5*(fcfc*Lcond/Condi$
;CONDENSER MOLECULAR FLOW PRESSURE DROP MODEL
* $Rhocm=((Pc+dPcc+dPcd)/(R*Tb))*M$
* $vcfcm = mdot/(Acf*Rhocm)$
* $Machcm = vcfcm/((9/7)*R*Tb/M)^.5$
* $Reccm = (Rhocm*vcfcm*Condid)/u$
* $dPcd = ((20+19*Lcond/Condr+3*(Lcond/Condr)^2)/(20+8*Lcond/Condr))*(mdot/(Condr^2$
* $fKnbc = FKnbc((Knc+Knbc)/2)$
* $dPcond = (1-fKnbc)*dPcnd+(fKnbc)*dPcd$
C $mf=(1-Knbc)/(Knc+Knbc)$
C $dPcond = mf*dPcnd+(1-mf)*dPcd$

;HEAT SHIELD CONTINUUM FLOW PRESSURE DROP MODEL

* $Rhonavhs=((Pc+dPcc+dPcond+dPhs)/(R*Tb))*M$
* $vcfhs = mdot/(Ahsa*Rhonavhs)$
* $Machhs = vcfhs/((9/7)*R*Tb/M)^.5$
* $Rechs = (Rhonavhs*vcfhs*Dhhs)/u$
* $fcfhs = 64/(Rechs*(8*Knhs*Fé+1))$
* $Alphahs = (2+8*Knhs*Fé+16*Knhs^2*Fé^2)/((4*Knhs*Fé+1)^2)$
* $PRhs = (Pc+dPcc+dPcond)/(Pc+dPcc+dPcond+dPhs)$
* $Alphahs/n*\ln(PRhs) = (n/(n+1))*((PRhs^{(1/n+1)-1}*R*Tb)/(M*vcfhs^2) + 0.5*(K2+K4)*ndP$
;BEFORE HEAT SHIELD CONTINUUM FLOW PRESSURE DROP MODEL

* $Rhonavbhs=((Pc+dPcc+dPcond+dPhs+dPbhs)/(R*Tb))*M$
* $vcfbhs = mdot/(Abhs*Rhonavbhs)$
* $Machbhs = vcfbhs/((9/7)*R*Tb/M)^.5$
* $Recbhs = (Rhonavbhs*vcfbhs*Cid)/u$
* $fcfbhs = 64/(Recbhs*(8*Knbs*Fé+1))$
* $Alphabhs = (2+8*Knbs*Fé+16*Knbs^2*Fé^2)/((4*Knbs*Fé+1)^2)$
* $PRbhs = (Pc+dPcc+dPcond+dPhs)/(Pc+dPcc+dPcond+dPhs+dPbhs)$
* $Alphabhs/n*\ln(PRbhs) = (n/(n+1))*((PRbhs^{(1/n+1)-1}*R*Tb)/(M*vcfbhs^2) + 0.5*(fcfbhs*($
;BASE TUBE/CELL WALL ANNULUS CONTINUUM FLOW PRESSURE DROP MODE
* $Rhonavan=((Pc+dPcc+dPcond+dPhs+dPbhs+dPan)/(R*Tb))*M$

S Rule

- * $vcfan = m\dot{o}/(\tilde{a}/4*(Cid^2-Dt^2)*Rhonavan)$
- * $Machan = vcfan/((9/7)*R*Tb/M)^{.5}$
- * $Recan = (Rhonavan*vcfan*Dhan)/u$
- * $ac=((Cid-Dt)^2*(Cid^2-Dt^2))/(Cid^4-Dt^4-((Cid^2-Dt^2)^2)/(\ln(Cid/Dt)))$
- * $fcfan = 64/(Recan*(8*Kne*Fé+1))$
- * $Alphaan = (2+8*Kne*Fé+16*Kne^2*Fé^2)/((4*Kne*Fé+1)^2)$
- * $PRan = (Pc+dPcc+dPcond+dPhs+dPbhs)/(Pc+dPcc+dPcond+dPhs+dPbhs+dPan)$
- * $Alphaan/n*\ln(PRan) = (n/(n+1))*((PRan^{(1/n+1)}-1)*R*Tb)/(M*vcfan^2) + 0.5*(fcfan*LB3/D$
;KNUDSEN NUMBER
- * $Kne=(2.596E-5*Tb/(Pc+dPcc+dPcond+dPhs+dPbhs+dPan))/((Cid-Dt)/2)$
- * $Kne1=u/Rhonavan*(\tilde{a}/2/R/Tb)^{0.5}/((Cid-Dt)/2)$
- * $Knbs=(2.596E-5*Tb/(Pc+dPcc+dPcond+dPhs+dPbhs))/Cid$
- * $Kne2=u/Rhonavbhs*(\tilde{a}/2/R/Tb)^{0.5}/Cid$
- * $Knhs=(2.596E-5*Tb/(Pc+dPcc+dPcond+dPhs))/Lhs$
- * $Kne3=u/Rhonavhs*(\tilde{a}/2/R/Tb)^{0.5}/Lhs$
- * $Knbc=(2.596E-5*Tb/(Pc+dPcc+dPcond))/Condid$
- * $Kne4=u/Rhonavc*(\tilde{a}/2/R/Tb)^{0.5}/Condid$
- * $Knc=(2.596E-5*Tb/(Pc+dPcc))/Condid$
- * $Kne5=u/Rhoc*(\tilde{a}/2/R/Tb)^{0.5}/Condid$
;DGM Model

<u>St</u>	<u>Input</u>	<u>Name</u>	<u>Output</u>	<u>Unit</u>	<u>Comment</u>
					RC1000 Pressure Loss Model
					INPUTS
	706.	Tb		C	BASE Avg. at Electrode 1
	718.	Tb2		C	BASE at Electrode 2
		dT	12.000	K	Secondary - Primary Temp Difference
	200.000	Tcond		C	Condenser Temperature
	8.000	I		A	Cell current
L	0.586	Voc2d		V	Electrode 2 Voltage from Data
L	0.500	Lhs		in	Distance from heat shield to plate
					PRESSURE RATIOS
		PRc	0.156		Condenser tube
		PRhs	0.974		Heat shield
		PRbhs	0.997		Between heat shield and BASE tube
		PRan	0.999		Annulus between cell wall and BASE t
					OPERATING CONDITIONS
		Voc2c	0.588	V	Corrected Voltage
	2.350E-4	Rcor		V/A	Resistance of Voltage Well
		dPerr	10.857		Pressure Loss Modeling Error
		j	0.218	A/cm^2	Tube Current density
					PRESSURE DROPS IN THE CELL
		Pc	0.018	Pa	Vapor pressure at condenser
		Pb	15076.170	Pa	Saturated vapor pressure at Tb
		Pb2	17558.993	Pa	Saturated vapor pressure at Tb2
L		Pe2c	17.925	Pa	Back Pressure at Electrode 2 from Da
L		Pe2	15.981	Pa	Back Pressure at Electrode 2
		dPcdd	17.907	Pa	Pressure Loss from Electrode 2 to Co
		dPcm	15.963	Pa	Pressure Loss from Electrode 2 to Co
L		dPan	0.019	Pa	BASE Tube/Cell Wall Annulus
L		dPbhs	0.041	Pa	Before Heat Shield
L		dPhs	0.418	Pa	Heat Shield/Cell Wall Annulus
L		dPcd	22.259	Pa	Condenser Molecular Flow
L		dPcnd	5.572	Pa	Condenser Continuum Flow
L		dPcond	14.469	Pa	Condenser
		dPcc	1.016	Pa	Condensing on condenser

ESTIMATING THE AREA OF ARTIFICIAL SPACE DEBRIS

Ronald A. Madler
Assistant Professor
Department of Aerospace Engineering

Embry-Riddle Aeronautical University
3200 Willow Creek Road
Prescott, Arizona 86301-3720

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

August 1996

ESTIMATING THE AREA OF ARTIFICIAL SPACE DEBRIS

Ronald A. Madler
Assistant Professor
Department of Aerospace Engineering
Embry-Riddle Aeronautical University

Abstract

The physical characteristics of breakup debris is essential to properly estimate the orbital debris environment and its hazard to spacecraft. This report presents, for the first time, a novel method to make direct measurements of the actual physical cross-sections of debris from ground-based breakups. The methodology of the laboratory experiments is presented and the results of these measurements are examined in comparison to other models for the physical characteristics of debris. For small objects, less than five cm in average diameter, the models and measurements agree well. The objects above five cm diverge from the predicted cross sectional area for debris. This may have a large effect on modeling of the debris environment [Madler (1994)].

ESTIMATING THE AREA OF ARTIFICIAL SPACE DEBRIS

Ronald A. Madler

Introduction

There exists a need to quantify and characterize the hazard posed to orbiting assets by space debris. Currently there are approximately 8000 objects being tracked by the Department of Defense (DoD) Space Surveillance Network (SSN). The majority of these tracked objects are operational and fragmentation debris. The tracked objects in orbit as of June 1996 can be categorized as either: active payloads (6%); inactive payloads (24%); rocket bodies (18%); operational debris (11%); anomalous event debris (2%); or satellite fragmentation debris (39%) [Grissom and Myers (1996)]. A much larger population of objects below the detection threshold is known to exist from the approximately 130 on-orbit satellite breakups which have occurred to date. Hence, it is important to be able to model these fragments from on-orbit fragmentation events. The relationships between mass, area, and size are very important because they affect the decay of space resident objects (RSOs), the expected collisional rate, and the hazard to other RSOs.

The mass-diameter and especially the mass-number relationship has been examined extensively [Bess (1975); Kessler and Cour-Palais (1978); Gravseth et al (1996, 1995, 1992), Maclay et al (1990, 1989)]. However, there has not been a concerted effort to measure the physical cross section of debris objects. This is mainly due to the difficulty in determining the cross section of the fragments. Several researchers have examined data from on-orbit fragmentations, but great uncertainties exist in those calculations [Anz-Meador et al (1995, 1993); Badhwar and Anz-Meador (1989); Dickey and Culp (1989)].

A novel method to determine the cross sectional area of debris fragments at a large number of orientations, from which an average value can be determined, has been developed. This method is both fast and effective. This study examines the relationships between the mass, area and size of fragments from ground-based tests.

Methodology/Discussion

Fragments from two ground-tests were obtained for his study, one a hypervelocity impact and one an explosion event. Shot CU-6470, a hypervelocity impact conducted at the Arnold Engineering Development Center (AEDC), was examined in cooperation with the University of Colorado and NASA/JSC, the owners of the fragments. The other set of fragments used came from an explosion test conducted by the European Space Agency/European Space Operations Center (ESA/ESOC). This explosion occurred on a downscaled ARIANE 4 H10 tank [Fucke (1993)].

The physical parameters size, mass and area were determined for a group of fragments from each event. Because the mass and size have been measured by many researchers, this study uses the same method for determining the characteristic dimension of fragments. Three orthogonal dimensions are determined from which the characteristic dimension is determined. The first dimension is the longest dimension of the object; the second dimension is the longest diameter perpendicular to the first; and the third dimension is the longest dimension perpendicular to the first two. The characteristic dimension is then the mean of these three dimensions. The mass and dimensions have been measured very precisely and accurately in this study. The dimensions have been measured using accurate calipers and are thus more precise than previous measurements.

The cross sectional area has not been determined through direct measurements before. This study uses a quick, direct method to determine the cross sectional area at a number of orientations. The method involved creating a uniform light field and placing the debris pieces in from of the light field to block a certain percentage of light. In essence it created a shadow graph. The view seen by the sensor is presented in Figure 28-1. The measured light level with and without the debris object in place can be used along with the known view area to determine the area of the debris object. This is represented by Equations 28-1 and 28-2. The distribution of areas and an average area are determined by taking measurements at 20 to 25 evenly spaced orientations.

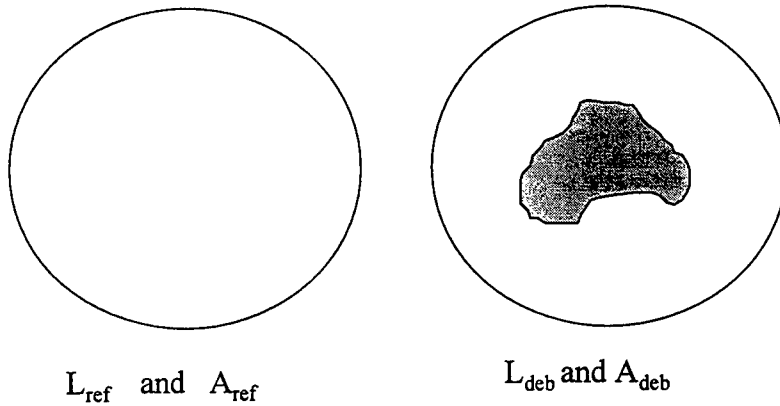


Figure 28-1: View down sensor tube as seen from sensor point of view

$$\frac{I * A_{ref}}{L_{ref}} = \frac{I * (A_{ref} - A_{deb})}{L_{deb}} \quad (28-1)$$

$$A_{deb} = A_{ref} * \left(1 - \frac{L_{deb}}{L_{ref}} \right) \quad (28-2)$$

where: I =light field intensity; A_{ref} =reference area; A_{deb} =debris area; L_{ref} =measured reference light;and L_{deb} =measured light with debris piece.

The measurements were taken at a special laboratory of the Laboratory for Atmospheric and Space Physics (LASP) of the University of Colorado, Boulder. The lab had a special black coating to the walls to prevent extraneous light from interfering with the measurements. The setup is shown in Figure 28-2 and explained in the following paragraphs.

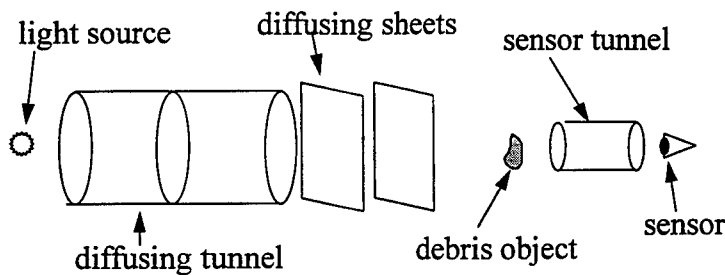


Figure 28-2: Laboratory setup.

The light source was a 500-Watt "Sun Gun" with a reflective backplate to direct light down the "diffusing tunnel." The diffusing tunnel was a roughened, white cardboard tube designed to produce an even light at the end. A plastic, photographic light diffuser was positioned midway down the four-foot, twelve inch diameter tube. At the end of the tube, two photographic mylar sheets provided the final diffusing elements to obtain an even light field which acts as a "light box," similar to ones used to view negatives. This setup was selected after many attempts to get a bright, even light field. measurements of the light field showed that it was uniform to within 5 or 10 percent, with the main variations occurring towards the edge of the light field. Due to this effect, the viewed region was limited to the inner portion of the light field.

The sensor measured the amount of light at the end of an eight inch, light-baffled tube. The sensor tube was baffled to prevent any forward scattering of light which could corrupt the data. Eye observations confirmed that little unwanted light was reaching the sensor. On the other end of the sensor tube, several cutouts of known dimension and area were placed to serve as a reference area. These different templates were used to minimize the reference area so that the ratio A_{deb}/A_{ref} was as high as possible to reduce error.

Measurements of two known objects, a black ping-pong ball and a floor hockey ball, were taken before each piece was examined over the 20 to 25 orientations. Since the area of these balls was well known, they can be used to verify/correct the area estimates. The debris objects were viewed at various orientations by suspending them with sewing thread and changing the orientation between each reading. Though a device was

created to aid in uniformly changing the orientation of the objects, it was found that having one person suspend the object at different orientations while another recorded the readings was more efficient. Multiple (usually two) runs were performed for each object. Due to limited lab time, 15 pieces from CU-6470 and 8 pieces from ESOC-2 were analyzed.

Results

The relative area frequency distribution for each individual object in the CU-6470 shot as well as a composite relative area distribution have been determined. The relative area is found by dividing each individual area estimate by the average area for a certain piece. The relative area distribution can then be compared for each piece regardless of the area differences. These distributions are compared with the relative area distributions calculated by measuring the shadows of several debris pieces. Ian Gravseth of the University of Colorado has contributed these measurements obtained using a planimeter. Figure 28-3 shows the relative area distribution for pieces 4 and 8 from CU-6470 calculated using the shadow method, while Figure 28-4 represents the planimeter data. While the planimeter data appears consistent between the two pieces, the shadow estimates do not. In fact one of the pieces appears to have a bi-modal area distribution. This was observed for many of the fragments. The composite relative area distribution was obtained by combining all the measurements for all 6470 objects, Figure 28-5. It is found to have a similar distribution as the planimeter data. More analysis is necessary to examine the area distribution of fragments.

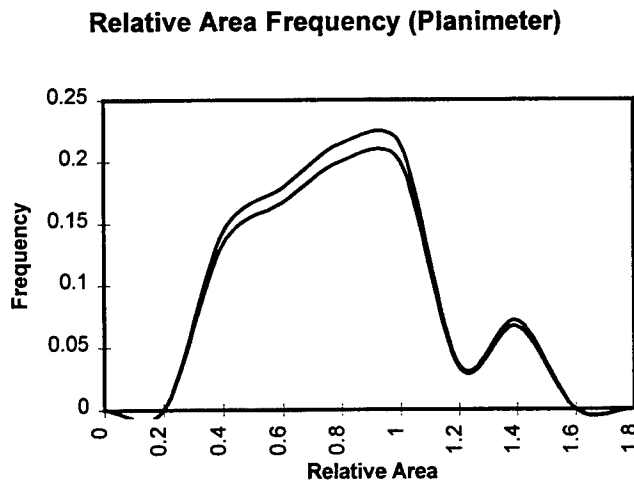


Figure 28-3: Relative area frequency distribution for pieces 4 and 8 from shot CU-6470. A slight bi-modal distribution is seen. This area was measured by a planimeter (courtesy of Ian Gravseth).

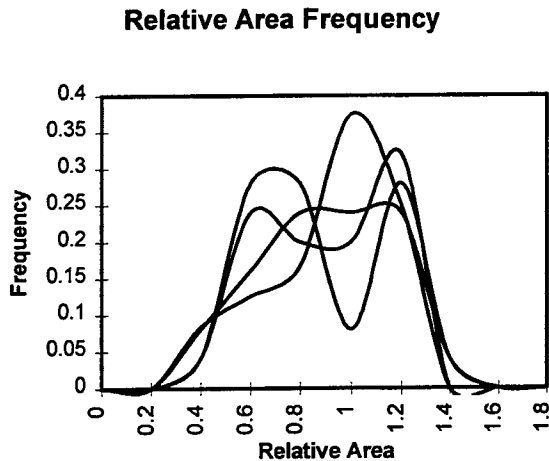


Figure 28-4: Shadow method for determining area for pieces 4 and 8 from shot CU-6470. The bi-modal distribution is apparent for one of the pieces.

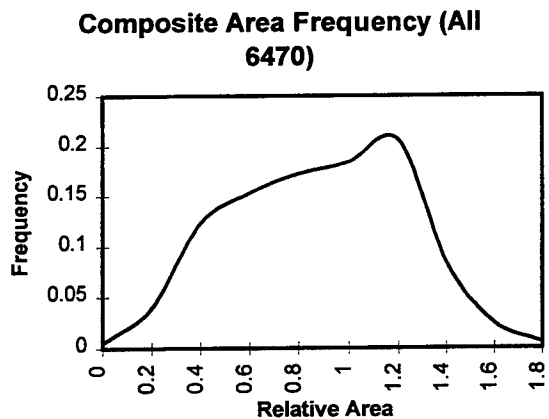


Figure 28-5: Composite relative area frequency for all shot CU-6470 pieces.

Figures 28-6 and 28-7 show the measured and estimated area versus mass for the two ground tests examined. The estimated area comes from the commonly used method of assuming the object is a sphere with the characteristic dimension as the diameter. It can be seen that the measured and estimated values diverge at larger masses. The mass range is significantly different between the two tests partly because the test objects had significantly different thicknesses. The diameter versus area comparisons follow a similar pattern as seen in Figures 28-8 and 28-9.

Area Versus Mass (CU-6470)

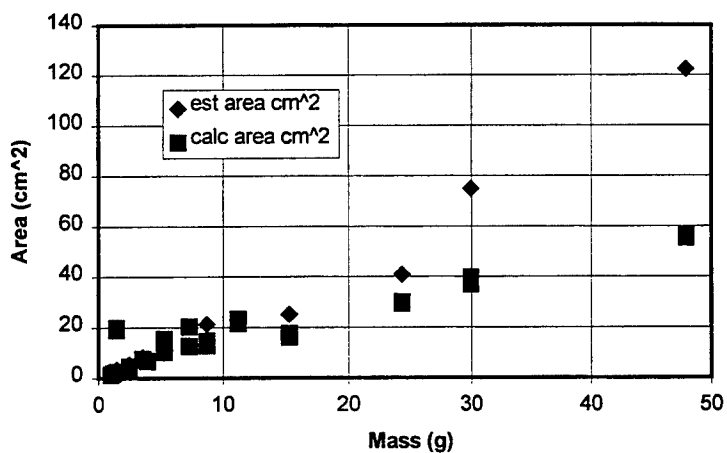


Figure 28-6: Area versus mass for CU-6470 objects. Estimated and measured values diverge at a certain point. There are two calculated (measured) values for each piece.

Area vs Mass (ESOC-2)

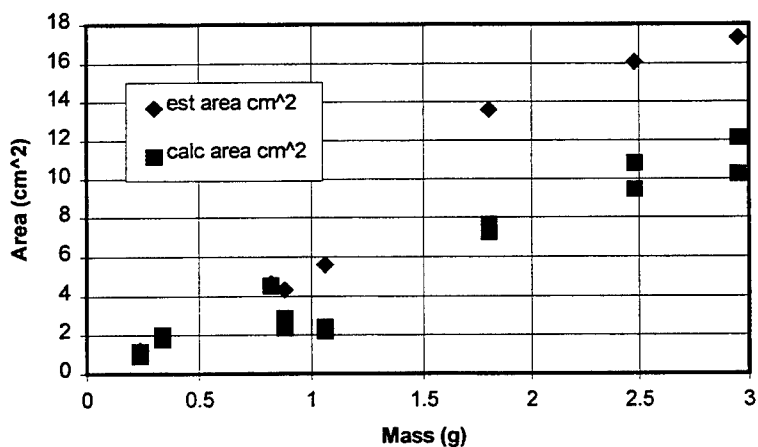


Figure 28-7: Area versus mass for ESOC-2 objects. Estimated and measured values diverge at a certain point. There are two calculated (measured) values for each piece.

Area Versus Mean Diameter (CU-6470)

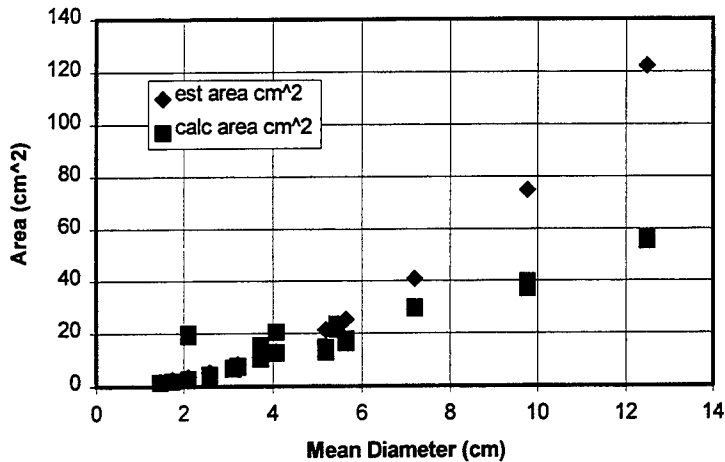


Figure 28-8: Area versus mean diameter for CU-6470 objects. Estimated and measured values diverge at a certain point. There are two calculated (measured) values for each piece.

Area vs Mean Diameter (ESOC-2)

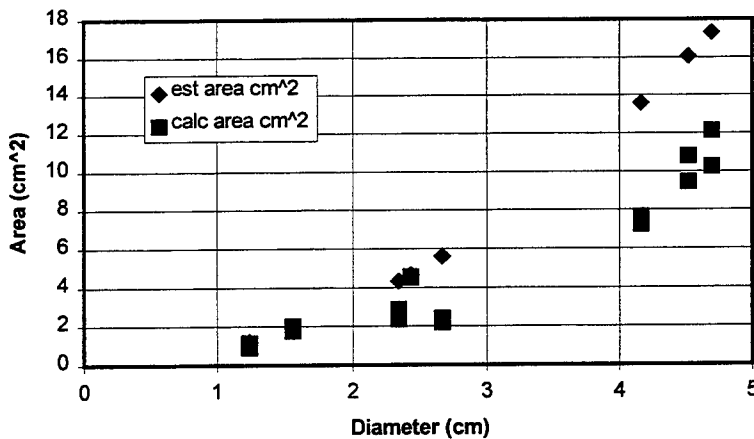


Figure 28-9: Area versus mean diameter for ESOC-2 objects. Estimated and measured values diverge at a certain point. There are two calculated (measured) values for each piece.

These results appear to imply that the usual method of estimating area may overestimate the area. This may have implications for debris models and measurements. The lower area would mean longer lifetimes within the models and hence a more severe future environment, while the lower area may mean that the measurements area actually seeing more massive objects than presently estimated. Both of these effects would make us believe

the environment is less hazardous than it actually is. Further investigation is necessary to clarify this point since this analysis is based on a small database.

The pieces analyzed in this report also appear to have a slightly different mass to diameter relationship than previously examined objects. While this is to be expected for the ESOC-2 data due to its smaller wall thickness, it is not known why the CU-6470 pieces appear to have a different relationship. Figures 28-10 and 28-11 show the area to mass relationships for the pieces and various model estimates.

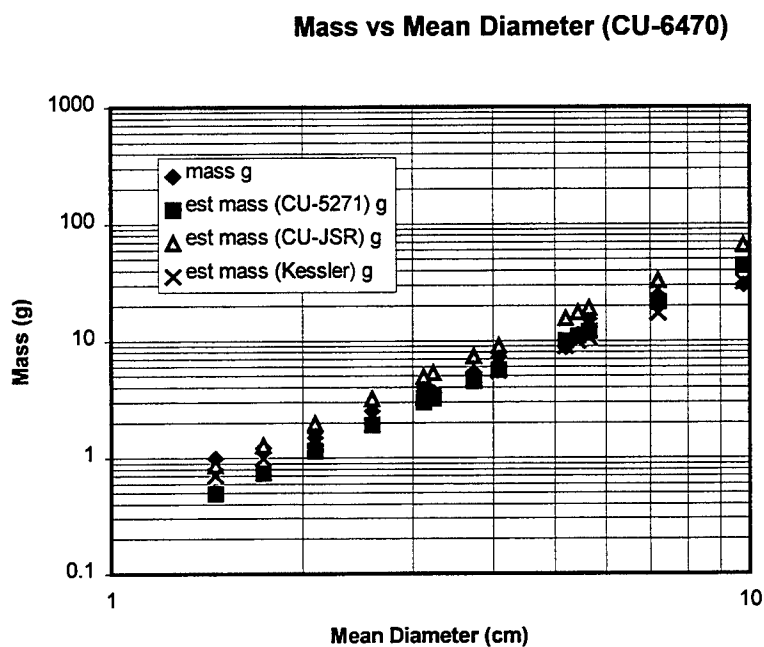


Figure 28-10: Measured and modeled mass versus mean diameter for CU-6470 objects. Three models are used to estimate the mass given a certain characteristic dimension. The estimated values appear close, but have a slightly different slope from the measured mass.

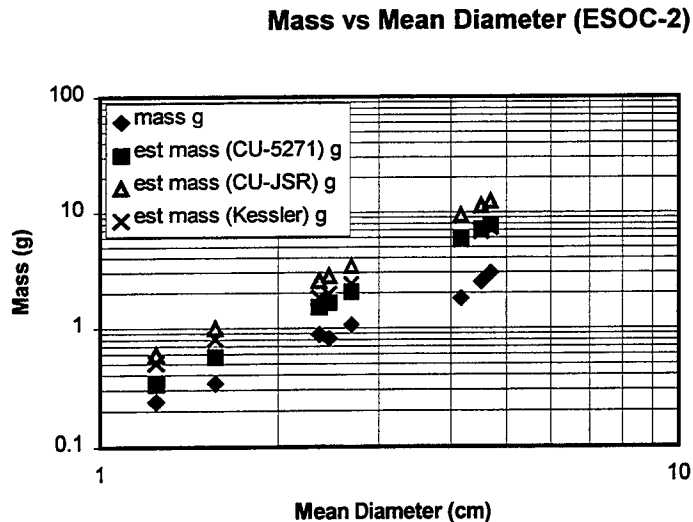


Figure 28-11: Measured and modeled mass versus mean diameter for ESOC-2 objects. Three models are used to estimate the mass given a certain characteristic dimension. The estimated values are way off because the ESOC-2 pieces have a very small thickness and hence low mass.

Conclusions/Summary

The purpose of this research has been to examine different methods for determining the cross sectional area of debris objects. For laboratory pieces, the mass, size, and area can be measured using several different techniques. The most straightforward method is to measure the projected area of the piece from several different orientations. This has been done using a planimeter (a device to measure area). This will give an area distribution as well as an average area, BUT is EXTREMELY slow and time consuming.

This research has developed a new, quicker approach for determining the cross sectional area of small debris objects. Objects are placed in front of a light source of known area. The intensity of the light is measured before and after the object is placed in front of the light source. The area of the object at that orientation can then be estimated. The main advantage of this method is that the area at each orientation can be determined very quickly and accurately.

The results appear to show that the actual cross sectional area is less than the estimated cross sectional area. This could have implications for our estimates of the debris environment. Further study is required to clarify this observation.

Future Work

Other approaches for determining the cross sectional area have been given a preliminary examination. A promising method which could be quick and does not require any specialized laboratory facilities uses digital imagery. Many images are taken of debris pieces from different orientations, and the images are analyzed to determine the cross sectional area. This method has only been tried on one debris piece so far, and the software was not completed during the summer research period. It appears to be a promising method for anyone interested in these statistics due to the emerging popularity of digital cameras. Figure 28-12 shows the image of a debris object and a reference object. Though the figure is black and white, the image is actually a color image. The color is important because the background is a unique color (red in this case) so that the image has high contrast when examined in the RGB sense. This allows the edges of the objects to be readily determined. As one can see from the image, it would be difficult to distinguish the edges in a black and white image.



Figure 28-12: A debris piece and a reference area imaged. Though in black and white it is hard to see the boundaries. The color image shows a red background, yellow ball, and dark debris object. Making it relatively easy to distinguish the edges of the objects.

Another proposed method is to use a vertical wind tunnel to suspend a rapidly spinning debris object in a stream of air. If the airspeed is known, it may be possible to determine a reasonably accurate cross section. This method may be examined in a follow-on study.

Future work is needed to determine a methodology to characterize the shape of debris objects. The shape of a fragment is necessary for determining the coefficient of drag (C_d) as well as the lethality of debris objects as a result of an on-orbit collision

Acknowledgments

The author would like to thank many people for their time, expertise, and equipment. In chronological order: Dr. David Spencer, Phillips Lab; AFOSR; Dr. Robert Culp, Dr. Dave DiLaura, Ian Gravseth, and Kira Jorgensen of the University of Colorado, Boulder; Drs. Bill McClintock, Rick Kohnert, and Scott Bailey of the Laboratory for Atmospheric and Space Physics (LASP); NASA/JSC for the use of their debris objects; and ESA for the use of their debris pieces. I would especially like to thank Kira Jorgensen, who was an able assistant during the laboratory testing of the debris pieces.

References

- Anz-Meador, P.D., Matney, M.J., and Kessler, D.J., "Physical Properties of Orbital Debris and Orbital Debris Clouds," *Advances in Space Research*, Vol. 16, No. 11, pp. 113-117, 1995.
- Anz-Meador, P.D., Rast, R.H., and Potter, A.E., Jr., "Apparent Densities of Orbital Debris," *Advances in Space Research*, Vol. 13, No. 8, pp. 153-156, 1993.
- Badhwar, G.D., and Anz-Meador, P.D., "Determination of the Area and Mass Distribution of Orbital Debris Fragments," *Earth, Moon, and Planets*, Vol. 45, pp. 29-51, 1989.
- Bess, T.D., "Size Distribution of Fragment Debris Produced by Simulated Meteoroid Impact of Spacecraft Wall," NASA SP-379, pp. 575-587, 1975.
- Dickey, M.R., and Culp, R.D., "Determining Characteristic Mass for Low-Earth-Orbiting Debris Objects," *Journal of Spacecraft and Rockets*, Vol. 26, No.6, pp. 460--464, 1989.
- Fucke, W., "Fragmentation Experiments for the Evaluation of the Small Size Debris Population," in *Proceedings of the First European Conference on Space Debris, Darmstadt, Germany, April 5-7, 1993*, W. Flury, Editor, ESA SD-01, pp. 275--280, July 1993.
- Gravseth, I.J., Maclay, T.D., and Culp, R.D., "Mass-Diameter and Characteristic-Length Ratio Functions for Orbital Debris," *Journal of Spacecraft and Rockets*, Vol. 33, No. 3, pp. 433-437, May-June 1996.

Gravseth, I.J., Maclay, T.D., and Culp, R.D., "A Mass-Diameter Function for the Entire Size Regime of Orbital Debris," in *Space Environmental, Legal, and Safety Issues*, Timothy D. Maclay, Editor, Proc. SPIE 2483, 21--30, 1995.

Gravseth, I., Maclay, T., "Analysis of Shot CU-5269 Fragments," Colorado Center for Astrodynamics Research Internal Report No. SD-92-02I, University of Colorado, August, 1992.

Gravseth, I., Maclay, T., "Analysis of Shot CU-5269 Fragments," Colorado Center for Astrodynamics Research Internal Report No. SD-92-01I, University of Colorado, February, 1992.

Kessler, D.J. and Cour-Palais, B.G., "Collision Frequency of Artificial Satellites: The Creation of a Debris Belt," *Journal of Geophysical Research*, Vol. 83, No. A6, June 1, 1978.

Maclay, T.D., and Hinga, M., "Analysis of Shot CU-5271 Fragments," Colorado Center for Astrodynamics Research Internal Report No. SD-90-01T, University of Colorado, March, 1990.

Maclay, T.D., and Hinga, M., "Analysis of Shot CU-5272 Fragments," Colorado Center for Astrodynamics Research Internal Report No. SD-90-02T, University of Colorado, June, 1990.

Maclay, T.D., Hinga, M., and Madler, R., "Analysis of Shot CU-6470 Fragments," Colorado Center for Astrodynamics Research Internal Report No. SD-89-02T, University of Colorado, July, 1989.

BOUNDARY CONDITIONS AT A PLASMA-FACING SURFACE

Carlos A. Ordonez
Assistant Professor
Department of Physics

University of North Texas
Denton, TX 76203-0368

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

August 1996

BOUNDARY CONDITIONS AT A PLASMA-FACING SURFACE

Carlos A. Ordóñez
Assistant Professor
Department of Physics
University of North Texas

Abstract

It has recently been reported that almost every commonly-used plasma-facing material has an electron emission coefficient, defined as the average number surface-emitted electrons per incident plasma electron, with a value larger than the minimum value at which space-charge saturation occurs for plasma temperatures above ~ 50 eV. With this motivation, a fully-kinetic self-consistent theory capable of describing the plasma sheath under conditions of space-charge saturation is developed. The theory is then used to obtain boundary conditions which are suitable for incorporation into computer programs which simulate plasmas. *More details of the work described here is contained in a paper submitted for publication to Physical Review E.*

BOUNDARY CONDITIONS AT A PLASMA-FACING SURFACE

Carlos A. Ordóñez

I. Introduction

Computer programs including magnetohydrodynamic (MHD), particle-in-cell (PIC), multi-fluid, and hybrid codes which simulate plasma processes are being increasingly incorporated in simulations of plasma-based devices. The types of plasma-based devices which can be (or have been) simulated by computer are diverse. Some examples are plasma-filled backward wave oscillators for producing high power microwaves,[1] plasma thrusters for space propulsion,[2] plasma wind tunnels,[3] opening switches including those of the compact toroid type,[4] thermionic converters,[5] plasma antennae,[6] "table-top" x-ray lasers,[6] electron-plasma-wave particle accelerators,[8,9] electron cyclotron resonance ion sources,[10,11] waste processors,[12] rail guns,[13] MHD generators,[14] plasma reactors for processing materials,[15,16] and fusion reactors.[17] In addition, reentry vehicles and spacecraft in low earth orbit should be mentioned.[18,19] Computer programs which simulate plasma processes often require boundary conditions at plasma/surface interfaces. In this paper, boundary conditions suitable for incorporation into such computer programs are presented which include the effect of the plasma sheath. A plasma sheath forms next to a plasma/surface interface and has a length scale (the Debye length) and a time scale (the inverse electron plasma frequency) which are often smaller than the length and time scales in computer simulations of macroscopic plasma processes. The plasma sheath is both non-neutral and non-Maxwellian and can have a substantial influence on particle and energy transport to a plasma-facing material surface. For example, in Ref. [20] the effect of a plasma sheath on particle and energy flow to an electrically-floating plasma-facing surface was reported. The plasma-facing surface bounded a simple magnetic mirror used to confine a collisional hydrogen-isotope plasma. The plasma sheath was found to be responsible for reducing the energy flow to

the plasma-facing surface by a factor of ten compared to that which would occur without the presence of the plasma sheath.

The boundary conditions developed in the work presented here are based on a fully-kinetic self-consistent theoretical description of the plasma sheath. Fully-kinetic sheath theory was initially developed by modeling the plasma-facing material surface as a particle sink.[21] The theory was later extended to include a surface emitting zero-temperature electrons.[22] Afterward, a surface emitting finite-temperature electrons was taken into account.[23] Fully-kinetic sheath theory has been developed by considering planar sources of ions and electrons and evaluating the resulting phase-space distributions for each. This type of fully-kinetic planar source approach has also been used for evaluating axial transport in the end region of a tandem mirror.[24-28]

Plasma near a solid material surface is normally divided into two regions. The region closest to the material surface has been referred to as the Debye sheath, the collector sheath, or simply, the sheath. The other region has been referred to as the presheath or the source sheath. The two regions together have also been referred to as the sheath. In the present work, the two regions are referred to as the presheath and the sheath. The boundary conditions presented are those that occur at the interface between the presheath and the sheath. (It should be mentioned that by incorporating the boundary conditions at the interface between the presheath and the sheath into a plasma simulation, the presheath region will naturally develop in the simulated plasma.) Depending on the plasma process considered, electrons and ions within the presheath can originate from a combination of mechanisms including plasma diffusion perpendicular and parallel to a magnetic field and ionization of gas atoms and molecules. A fundamental assumption in the fully-kinetic planar source approach to sheath theory is that all presheaths can be adequately modeled as a collisionless plasma region which is bounded on one side by a planar source of half-Maxwellian electrons and ions and on the other side by

the presheath/sheath interface.[21] This fundamental assumption allows for the plasma description both at the presheath/sheath interface and at any location within the sheath to be self-consistently determined.

Most of the commonly-used plasma-facing surface materials have recently been found to have secondary electron emission coefficients near or above 0.9 at moderate edge-plasma temperatures.[29] (An edge plasma is defined at present as a Maxwellian plasma which supplies ions and electrons to a presheath.) For example, the following materials have secondary electron emission coefficients which reach 0.9 at the edge-plasma temperatures indicated (these values are from Fig. 3 of Ref. [29]): boron, 15 eV; carbon, 42 eV; aluminum, 47 eV; silicon, 29 eV; titanium, 49 eV; iron, 35 eV; copper, 52 eV; molybdenum, 35 eV; and tungsten, 53 eV. The edge-plasma temperature at which the secondary electron emission coefficient reaches ~ 0.9 is important since the onset of space-charge saturation has been found to take place within the plasma sheath when the secondary electron emission coefficient is ~ 0.9 . [23] In Ref. [23], the sheath description is determined self-consistently for a monotonically decreasing potential including the onset of space-charge saturation when the electric field is zero at the wall surface. Ref. [23] provides a historical review of sheath theory and a detailed comparison of the fully-kinetic approach with other approaches to sheath theory. In the present work, the sheath theory in Ref. [23] is extended by self-consistently determining the sheath properties not only for a monotonically decreasing potential but also for space-charge saturation when the electric field is reversed near the wall surface and a potential minimum occurs within the sheath.

One aspect of fully-kinetic sheath theory, which makes it difficult to use, is that evaluation of the sheath and presheath potentials requires a numerical solution to a set of coupled non-linear equations. Although boundary conditions can be evaluated in closed form, they are in terms of the sheath and presheath potentials. In order to provide

for fast computation of the boundary conditions, simple expressions are developed for the sheath and presheath potentials which incorporate fits to numerically calculated values. The problem is considered in a four-dimensional phase space which consists of one spacial and three velocity dimensions with the spacial dimension normal to the plane of the plasma-facing surface. The two velocity dimensions perpendicular to the spacial dimension are symmetric and the theory can, in principle, be reduced from three to two velocity dimensions without loss of information.

The properties of a plasma sheath depend upon a number of things including plasma composition, ion charge-state distribution, magnetic field angle with respect to the surface normal, plasma flow speed parallel to the surface, and charged-particle emission processes at the surface. The emission processes are affected by such things as surface: temperature, composition, cleanness, roughness, electric fields and magnetic fields. The plasma-facing material surface can be either a current-carrying anode or cathode, or it can electrically float with respect to the plasma. Altogether, a complete theoretical description of a plasma sheath is extremely difficult to achieve and some simplifying approximations are necessary. In the present work, all ions are considered to have the same mass and charge state. The effect of a magnetic field is not taken into account and no plasma flow parallel to the material surface is considered. The sheath is collisionless and bounded at a planar sheath/surface interface by an electrically-floating wall which absorbs all incident particles and emits only secondary and thermionic electrons.

In Section II, the presheath and sheath potentials are evaluated numerically. In Section III, the boundary conditions at the presheath/sheath interface are given along with the simple expressions for the presheath and sheath potentials. A concluding summary is found in Section IV.

II. Evaluation of the Sheath and Presheath Potentials

A number of conditions are implemented in order to evaluate the sheath and presheath potentials. The first two conditions are that the total current density to the electrically-floating surface is zero (for a steady-state solution) and that the charge density at the presheath/sheath interface is zero. The third condition is that the spacial integral of the charge density in the presheath is zero. This requires the presheath to be globally quasineutral. It should be noted, however, that a manifestation of the presheath model used here is that the presheath is not locally quasineutral except at the presheath/sheath interface. The fourth condition, which applies only when space-charge saturation takes place within the sheath, is that the electric field at the electric potential minimum is zero.

The first condition, zero current density, requires $Z F_i = F_e - F_\delta$. This condition provides the relations,

$$\alpha = e^{-\psi_{spe}} (1 - \delta) \sqrt{\frac{\eta}{\tau_i}} \quad (1)$$

$$\alpha = e^{-\psi_{mpe}} (1 - \delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} \quad (2)$$

and

$$\alpha = e^{\psi_{spi} - \psi_{mpe}} (1 - \delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} \quad (3)$$

which correspond to the potential profiles shown in Figs. 1a, 1b, and 1c, respectively. Here, $\alpha = Z n_{pi}/n_{pe}$ is a parameter called the neutralization factor,[21] $\eta = m_i/m_e$, and $\tau_i = T_{pi}/T_{pe}$.

In order to implement the second and third conditions, the charge density, $\rho = e(Z n_i - n_e - n_\delta)$, between x_p and x_b is needed. The charge density can be written as a set of three relations in terms of α corresponding to the three potential profiles.

Substituting for α provides,

$$\frac{2\rho e^{\psi_{spe}}}{en_{pe}} = (1-\delta) \sqrt{\frac{\eta}{\tau_i}} G_1(\psi_{pxi}) - G_2(\psi_{sxe}) - \frac{\delta}{\sqrt{\tau_\delta}} G_1(\psi_{sxd}) \quad (4)$$

$$\frac{2\rho e^{\psi_{mpe}}}{en_{pe}} = (1-\delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} G_1(\psi_{pxi}) - G_2(\psi_{mxe}) - \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} G_1(\psi_{mxd}) \quad (5)$$

$$\begin{aligned} \frac{2\rho e^{\psi_{mpe}}}{en_{pe}} = & (1-\delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} [2e^{\psi_{spi}} G_1(\psi_{pxi}) - G_1(\psi_{sxi})] \\ & - G_2(\psi_{mxe}) - \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} G_1(\psi_{mxd}) \end{aligned} \quad (6)$$

where $\tau_\delta = T_{s\delta}/T_{pe}$. The second condition, zero charge density at the presheath/sheath interface, requires the right-hand-sides of the above relations to equal zero at x_b . This written as

$$(1-\delta) \sqrt{\frac{\eta}{\tau_i}} G_1(\psi_{pbi}) = G_2(\psi_{sbe}) + \frac{\delta}{\sqrt{\tau_\delta}} G_1(\psi_{sbd}) \quad (7)$$

$$(1-\delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} G_1(\psi_{pbi}) = G_2(\psi_{mbe}) + \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} G_1(\psi_{mbd}) \quad (8)$$

and

$$(1-\delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} [2e^{\psi_{spi}} G_1(\psi_{pbi}) - G_1(\psi_{sbi})] = G_2(\psi_{mbe}) + \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} G_1(\psi_{mbd}) \quad (9)$$

The third condition, that the spacial integral of the charge density in the presheath is zero, requires zero electric fields at x_p and x_b . With no electric fields at the edge-plasma/presheath interface and at the presheath/sheath interface, it can be shown (see, for example, Eq. (36) of Ref. [22]) that $\int_{\phi_p}^{\phi_b} \rho d\phi = 0$ or, equivalently, $\int_{\psi_{pe}}^{\psi_{be}} \rho d\psi_{xe} = 0$. Consequently, the third condition requires

$$(1-\delta) \sqrt{\frac{\eta}{\tau_i}} \int_{\psi_{pe}}^{\psi_{be}} G_1(\psi_{pxi}) d\psi_{xe} = \int_{\psi_{pe}}^{\psi_{be}} G_2(\psi_{sxe}) d\psi_{xe} + \frac{\delta}{\sqrt{\tau_\delta}} \int_{\psi_{pe}}^{\psi_{be}} G_1(\psi_{sxd}) d\psi_{xe} \quad (10)$$

$$(1-\delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} \int_{\psi_{pe}}^{\psi_{be}} G_1(\psi_{pxi}) d\psi_{xe}$$

$$= \int_{\psi_{pe}}^{\psi_{be}} G_2(\psi_{mxe}) d\psi_{xe} + \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} \int_{\psi_{pe}}^{\psi_{be}} G_1(\psi_{mx\delta}) d\psi_{xe} \quad (11)$$

$$(1 - \delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} \left[2 e^{\psi_{spi}} \int_{\psi_{pe}}^{\psi_{be}} G_1(\psi_{pxi}) d\psi_{xe} - \int_{\psi_{pe}}^{\psi_{be}} G_1(\psi_{sxi}) d\psi_{xe} \right]$$

$$= \int_{\psi_{pe}}^{\psi_{be}} G_2(\psi_{mxe}) d\psi_{xe} + \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} \int_{\psi_{pe}}^{\psi_{be}} G_1(\psi_{mx\delta}) d\psi_{xe} \quad (12)$$

Since the electric field is zero at the presheath/sheath interface, the fourth condition, which requires the electric field at the electric potential minimum to be zero, is equivalent to requiring the spacial integral of the charge density between the presheath/sheath interface and the potential minimum to be zero. Hence, the fourth condition is similar to the third except that it applies only to the profiles shown in Figs. 1b and 1c and that it applies between x_b and x_m . The fourth condition requires,

$$(1 - \delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} \int_{\psi_{be}}^{\psi_{me}} G_1(\psi_{pxi}) d\psi_{xe}$$

$$= \int_{\psi_{be}}^{\psi_{me}} G_2(\psi_{mxe}) d\psi_{xe} + \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} \int_{\psi_{be}}^{\psi_{me}} G_1(\psi_{mx\delta}) d\psi_{xe} \quad (13)$$

$$(1 - \delta e^{-\psi_{ms\delta}}) \sqrt{\frac{\eta}{\tau_i}} \left[2 e^{\psi_{spi}} \int_{\psi_{be}}^{\psi_{me}} G_1(\psi_{pxi}) d\psi_{xe} - \int_{\psi_{be}}^{\psi_{me}} G_1(\psi_{sxi}) d\psi_{xe} \right]$$

$$= \int_{\psi_{be}}^{\psi_{me}} G_2(\psi_{mxe}) d\psi_{xe} + \frac{\delta}{\sqrt{\tau_\delta}} e^{-\psi_{ms\delta}} \int_{\psi_{be}}^{\psi_{me}} G_1(\psi_{mx\delta}) d\psi_{xe} \quad (14)$$

The above integrals are given by the two indefinite integrals,

$$\int G_1(x) dx = G_1(x) + 2 \sqrt{\frac{x}{\pi}} \quad (15)$$

and

$$\int G_2(x) dx = G_2(x) - 2 \sqrt{\frac{x}{\pi}} \quad (16)$$

along with $\psi_i = -Z\psi_e/\tau_i$ and $\psi_\delta = \psi_e/\tau_\delta$. The electric potential at one location must also be defined. Hereafter, the electric potential at the presheath/sheath interface shall be defined as zero, $\psi_{be} = 0$. With $\psi_{be} = 0$, ψ_{se} gives the normalized sheath potential drop (as a positive value for a monotonically decreasing potential) and ψ_{pe} gives the normalized presheath potential drop (as a negative value). For a monotonically decreasing potential (Fig. 1a), the associated equations are solved simultaneously for the two unknown parameters, ψ_{pe} and ψ_{se} . For a space-charge saturated sheath, the three unknown parameters, ψ_{pe} , ψ_{me} and ψ_{se} , are solved simultaneously using the associated equations. In order to determine the transition from the potential profile in Fig. 1a to that in Fig. 1b (this transition occurs at the onset of space-charge saturation), the needed equations are solved simultaneously for δ , ψ_{pe} and ψ_{se} under the condition, $\psi_{me} = \psi_{se}$. In order to determine the transition from the potential profile in Fig. 1b to that in Fig. 1c, the needed equations are solved simultaneously for δ , ψ_{pe} and ψ_{me} under the condition, $\psi_{se} = \psi_{pe}$.

Once the values for the normalized potentials are known, the profile of the electric potential within the presheath and sheath is evaluated using Poisson's equation. The procedure is outlined in Ref. [22]. Eq. (39) of Ref. [22] is written here as

$$\frac{x}{\lambda_{D(p)}} = \int_{\psi_{pe}}^{\psi_{xe}} \left[\int_{\psi_{pe}}^{\psi'_{xe}} \frac{2\rho(\psi''_{xe})}{e n_{pe}} d\psi''_{xe} \right]^{-1/2} d\psi'_{xe} \quad (x_p < x < x_m) \quad (17)$$

where $x_p = 0$, $\lambda_{D(p)}$ is the Debye length in the edge plasma. The inner integral of this relation is evaluated in closed form while the outer is evaluated numerically. The results are then inverted to obtain ψ_{xe} as a function of x . Fig. 1 shows actual self-consistent calculations of different electric potential profiles. (The vertical dimensions are different for each of the three profiles shown in Fig. 1. The horizontal dimensions are the same, however, and the width of the curve in Fig. 1a is $27 \lambda_{D(p)}$.) The profiles are calculated for a thermal ($\tau_i = 1$) hydrogen ($Z = 1, \eta = 1836$) plasma which is bounded by an electron-emitting surface (with $\tau_\delta = 0.2$). The profiles are shown in order of increasing

values for the electron emission coefficient: $\delta = 0$ for Fig. 1a, $\delta = 15$ for Fig. 1b, and $\delta = 1000$ for Fig. 1c. The values for the normalized potentials calculated for each of the profiles are: $\psi_{pe} = -0.34$ and $\psi_{se} = 2.5$ for Fig. 1a; $\psi_{pe} = -0.43$, $\psi_{me} = 0.29$ and $\psi_{se} = -0.28$ for Fig. 1b; and, $\psi_{pe} = -0.53$, $\psi_{me} = 0.31$ and $\psi_{se} = -1.1$ for Fig. 1c. The onset of space-charge saturation occurs at $\delta = 0.88$, $\psi_{pe} = -0.43$, and $\psi_{me} = \psi_{se} = 0.29$ while the transition from a negative to a positive surface floating potential occurs at $\delta = 31$, $\psi_{pe} = \psi_{se} = -0.43$, and $\psi_{me} = 0.29$.

III. Boundary Conditions

The boundary conditions provided in this section are those which occur at the presheath/sheath interface. This location is both quasineutral and electric-field free. For brevity, location subscripts are not used in this section since quantities which were previously location dependent are now only considered at the presheath/sheath interface. The boundary conditions at the presheath/sheath interface are the pressure terms, $\langle v_y^2 \rangle$, $\langle v_z^2 \rangle$ and $\langle v_x^2 \rangle - \langle v_x \rangle^2$, the plasma flow velocity, $\langle v_x \rangle$, and the energy flow term, $\langle v^2 v_x \rangle$. Using relations from Section III, the boundary conditions are written in terms of the electron and ion plasma temperatures (T_e and T_i) at the presheath/sheath interface, the temperature (T_δ) associated with electrons emitted at the plasma-facing surface, the ratio (τ_i) of ion to electron plasma temperatures at the edge plasma, the normalized presheath (ψ_{pe}), sheath (ψ_{se}), and minimum (ψ_{me}) potentials, the electron and ion masses (m_e and m_i) and the ion charge state (Z). Using relations in Section IV, the normalized potentials are evaluated numerically and the values obtained are fit in terms of Z , τ_i , the electron emission coefficient (δ), which equals the ratio of surface-emitted electron flux to plasma electron flux incident on the surface, the ratio (η) of ion to electron mass, and the ratio (τ_δ) of the temperature associated with electrons emitted at the surface to the electron plasma temperature at the edge plasma. Of the three types of electric potential profiles shown in Fig. 1, the third is expected to be rare since, for $\tau_\delta \leq 0.1$, the electron emission coefficient must be very large, $\delta > 10^3$, to produce the type of potential

profile shown in Fig. 1c (with $\tau_i \sim 1$). For this reason, only boundary conditions for the electric potential profiles shown in Figs. 1a and 1b are presented. Also, all calculations of normalized potentials are carried out using the same τ_δ value, $\tau_\delta = 0.01$. This is appropriate since, for $\tau_\delta \leq 0.1$, the normalized potentials are essentially independent of τ_δ .

For use of the boundary conditions, τ_i is the only non-local parameter needed. The value for τ_i should be evaluated at one of two possible locations depending on which of the two has a shorter path length to the wall surface (the path length should be along a magnetic field line if a magnetic field is present). One location is an ion mean-free-path away from the wall surface while the other location is where the maximum temperature occurs along the magnetic field line. The first of these two locations applies when no magnetic field is present or when the source of presheath ions is predominantly either ionization or diffusion parallel to a magnetic field; the second method applies when diffusion perpendicular to a magnetic field provides most of the presheath ions.

Two of the boundary conditions, the second and third pressure terms, can be written in terms of another boundary condition, the first pressure term. They are $\langle v_z^2 \rangle = \langle v_y^2 \rangle$ and

$$\langle v_x^2 \rangle - \langle v_x \rangle^2 = \frac{3T}{m} - 2 \langle v_y^2 \rangle \quad (18)$$

The other boundary conditions are determined separately for plasma electrons, plasma ions, and surface-emitted electrons. (For surface-emitted electrons, $\langle v_x \rangle_\delta$ and $\langle v^2 v_x \rangle_\delta$ are defined as positive in the negative-x direction.) For the electric potential profile shown in Fig. 1a, the boundary conditions are

$$\langle v_y^2 \rangle_e = \frac{T_e}{m_e \chi_{se}} \quad (19)$$

$$\langle v_x \rangle_e = \frac{\sqrt{2 \langle v_y^2 \rangle_e / \pi}}{e^{\psi_{se}} \operatorname{erfc}(-\sqrt{\psi_{se}})} \quad (20)$$

$$\langle v^2 v_x \rangle_e = 2 (2 + \psi_{se}) \langle v_y^2 \rangle_e \langle v_x \rangle_e \quad (21)$$

$$\langle v_y^2 \rangle_i = \frac{T_i}{m_i \chi_{pi}} \quad (22)$$

$$\langle v_x \rangle_i = \frac{\sqrt{2 \langle v_y^2 \rangle_i / \pi}}{e^{\psi_{pi}} \operatorname{erfc}(\sqrt{\psi_{pi}})} \quad (23)$$

$$\langle v^2 v_x \rangle_i = 2 (2 + \psi_{pi}) \langle v_y^2 \rangle_i \langle v_x \rangle_i \quad (24)$$

$$\langle v_y^2 \rangle_\delta = \frac{T_\delta}{m_e} \quad (25)$$

$$\langle v_x \rangle_\delta = \frac{\sqrt{2 \langle v_y^2 \rangle_\delta / \pi}}{e^{\psi_{s\delta}} \operatorname{erfc}(\sqrt{\psi_{s\delta}})} \quad (26)$$

and

$$\langle v^2 v_x \rangle_\delta = 2 (2 + \psi_{s\delta}) \langle v_y^2 \rangle_\delta \langle v_x \rangle_\delta \quad (27)$$

where the following definitions are used:

$$\chi_{se} = 1 - \frac{2 \sqrt{\psi_{se}/\pi}}{3 e^{\psi_{se}} \operatorname{erfc}(-\sqrt{\psi_{se}})} - \frac{2}{3 \pi [e^{\psi_{se}} \operatorname{erfc}(-\sqrt{\psi_{se}})]^2} \approx \ln (2.51 \psi_{se}^{0.042}) \quad (28)$$

$$\chi_{pi} = 1 + \frac{2 \sqrt{\psi_{pi}/\pi}}{3 e^{\psi_{pi}} \operatorname{erfc}(\sqrt{\psi_{pi}})} - \frac{2}{3 \pi [e^{\psi_{pi}} \operatorname{erfc}(\sqrt{\psi_{pi}})]^2} \approx \ln (2.05 \psi_{pi}^{-0.013}) \quad (29)$$

$\psi_{pi} = -Z \psi_{pe}/\tau_i$, and $\psi_{s\delta} = T_e \psi_{se}/(T_\delta \chi_{se})$. The approximations for χ_{se} and χ_{pi} are accurate to within 2% for $0.1 < \psi_{se} < 10$ and $0.01 < \psi_{pi} < 100$, respectively. For the normalized sheath and presheath potentials, the following two expressions are recommended:

$$\psi_{se} = \ln [0.2725 (1 - \delta) \sqrt{\eta} \tau_i^{-0.092} Z^{-0.44}] \quad (30)$$

$$\psi_{pe} = -\ln \left[1.43 \left(\frac{Z}{\tau_i} \right)^{0.195} \right] \quad (31)$$

These expressions are fits to numerically-determined values using the values: 1, 10, and 100 times the proton-to-electron mass ratio (1836) for η ; 1/3, 1, and 3 for τ_i ; 1, 2, and 3 for Z ; and, 0 and 1/2 for δ . The numerically-determined values were calculated using all possible combinations of η , τ_i , Z , and δ except combinations which simultaneously

involved $Z = 1$ and $\eta = 1836$. The expressions for ψ_{se} and ψ_{pe} agree to within 4% and 7%, respectively, with the numerically-determined values. Notice that the relation for the sheath potential has the usual, $\ln[(1 - \delta) \sqrt{\eta}]$, dependence given elsewhere.[30,31]

The onset of space-charge saturation and the corresponding transition from a monotonically decreasing potential (e.g. Fig. 1a) to a potential profile with a potential minimum (e.g. Fig. 1b) takes place if the electron emission coefficient reaches a sufficiently large value. A suitable expression which provides the value of the electron emission coefficient at the transition between profiles is

$$\delta_c = 1 - 7.7 \tau_i^{0.1} \sqrt{\frac{Z}{\eta}} \quad (32)$$

This expression is a fit to values determined numerically for the same combinations of η , τ_i , and Z which were used for the ψ_{se} and ψ_{pe} fits. The δ_c fit is remarkably accurate being within 0.1% of the numerically-determined values.

If the electron emission coefficient is larger than δ_c , boundary conditions associated with the potential profile shown in Fig. 1b should be used. For the potential profile shown in Fig. 1b, the boundary conditions are the same as those for Fig. 1a except with subscript, s , replaced by subscript, m . For the normalized sheath minimum and presheath potentials under space-charge saturation conditions (Fig. 1b), the following two expressions are recommended:

$$\psi_{me} = \ln(1.63 \eta^{0.01}) \quad (33)$$

and

$$\psi_{pe} = -\ln \left[1.51 \left(\frac{Z}{\tau_i} \right)^{0.209} \right] \quad (34)$$

These two expressions agree to within 5% and 6%, respectively, with numerically-determined values using the combinations of η , τ_i , and Z used for the δ_c fit along with $\delta = \delta_c$ and

$\delta = 10$. Note that these expressions are independent of δ for $\delta \geq \delta_c$. Thus, a specific relation for the electron emission coefficient under conditions of space-charge saturation is not needed for calculating the associated boundary conditions. It should also be mentioned that trapping of slow charge-exchange ions within the potential well surrounding the potential minimum has been found to be responsible for reducing space-charge saturation.[32] When that is the case, it is probable that the sheath potential profile is adequately described by using an effective electron emission coefficient which is smaller than the actual one. If so, then for situations in which $\delta > \delta_c$, the boundary conditions presented here are unaffected by ion trapping since the sheath minimum and presheath potentials are not dependent on δ . Nevertheless, a more detailed study of the effects of ion trapping on space-charge saturation appears warranted.

IV. Concluding Summary

A fully-kinetic self-consistent theory describing the plasma sheath under conditions of space-charge saturation has been presented. The phase-space distribution functions for each species of particles have been evaluated and velocity moments have been taken in order to obtain particle densities, fluxes, temperatures, and energy fluxes. The electric potential profile has been determined self-consistently for three different types of profiles. These are a monotonically-decreasing potential, a single-minimum potential associated with a negative surface floating potential, and a single-minimum potential associated with a positive surface floating potential. The electron emission requirements for production of the third potential profile were found to be severe. For this reason, boundary conditions have been provided only for the first two of the three potential profiles. Boundary conditions on velocity moments have been supplied for three separate species: plasma electrons, plasma ions, and surface-emitted electrons. These boundary conditions can be combined as needed for use in multi-fluid and single-fluid (MHD) applications. The boundary conditions are in terms of sheath and presheath potential drops and simple expressions have been provided for these potential drops. Once the

potential drop expressions are inserted into the boundary conditions, the boundary conditions are in terms of the three species' temperatures, τ_i , the ion-to-electron plasma temperature ratio, Z , the ion charge state, η , the ion-to-electron mass ratio, and δ , the electron emission coefficient.

In order to determine the value of δ at which the transition between the first two potential profiles occurs, the following expression was developed: $\delta_c = 1 - 7.7 \tau_i^{0.1} \sqrt{Z/\eta}$. This expression gives the value of the electron emission coefficient at the onset of space-charge saturation. For $\delta < \delta_c$, the following expressions are recommended for the sheath and presheath potentials normalized to the electron-plasma temperature: $\ln[0.2725 (1 - \delta) \sqrt{\eta} \tau_i^{-0.092} Z^{-0.44}]$ and $\ln[1.43 (Z/\tau_i)^{0.195}]$. For $\delta \geq \delta_c$, recommended expressions for the normalized sheath minimum and presheath potentials are, respectively, $\ln(1.63 \eta^{0.01})$ and $\ln[1.51 (Z/\tau_i)^{0.209}]$. The conditions these expressions for δ_c and the normalized potential are expected to be suitable for are: a temperature associated with the emitted electrons which is less than one tenth that associated with the plasma electrons; $1/3 \leq \tau_i \leq 3$; $1 \leq Z \leq 3$; and $1836 \leq \eta \leq 100 \times 1836$.

References

- [1]X. Zhai, E. Garate, R. Prohaska, A. Fisher, and G. Benford, *Phys. Lett. A* **186**, 330 (1994).
- [2]P. J. Turchi in *Space Propulsion Analysis and Design*, edited by R. W. Humble, G. N. Henry, and W. J. Larson, (McGraw-Hill, Inc., New York, 1995) p. 509.
- [3]S. Fasoulas, P. C. Sleazona, M. Auweter-Kurtz, H. A. Habiger, S. H. Laure, and A. T. Schonemann, *J. Thermophysics and Heat Transfer* **9**, 422 (1995).
- [4]R. E. Peterkin, Jr., J. H. Degnan, T. W. Hussey, N. F. Roderick, and P. J. Turchi, *IEEE Trans. Plasma Sci.* **21**, 522 (1993).
- [5]M. L. Ramalingam, "Advanced Thermionic Technology Initiative Program," U. S. Air Force Wright Lab., WL-TR-92-2004, Dayton, OH, 1992.
- [6]B. D. Nordwall, *Aviation Week & Space Technol.*, June 10, 50, 1996.
- [7]T. Wagner, E. Eberl, K. Frank, W. Hartmann, D. H. H. Hoffmann, and R. Tkotz, *Phys. Rev. Lett.* **76**, 3124 (1996).
- [8]J. R. Marques, J. P. Geindre, F. Amiranoff, P. Audebert, J. C. Gauthier, A. Antonetti, and G. Grillon, *Phys. Rev. Lett.* **76**, 3566 (1996).
- [9]C. W. Siders, S. P. Le Blanc, D. Fisher, T. Tajima, M. C. Downer, A. Babine, A. Stepanov, and A. Sergeev, *Phys. Rev. Lett.* **76**, 3570 (1996).

- [10]P. Sortais, Rev. Sci. Instrum. **67**, 867 (1996).
- [11]M. Sekiguchi, Rev. Sci. Instrum. **67**, 1606 (1996).
- [12]G. Ondrey and K. Fouhy, Chem. Engineering **98** (12) 32 (1991).
- [13]Proceedings of the 7th Symposium on Electromagnetic Launch Technology, IEEE Trans. Magnetics **31** (1995).
- [14]V. R. Malghan, Energy Convers. Mgmt. **37** 569 (1996).
- [15]M. Kenward, Phys. World **8** (6) 31 (1995).
- [16]D. B. Graves, IEEE Trans. Plasma Sci. **22**, 31 (1994).
- [17]J. Sheffield, Rev. Modern Phys. **66** 1015 (1994).
- [18]J. V. Shebalin, J. Spacecraft **28** 394 (1991).
- [19]D. E. Hastings, J. Geophys. Res. **100** 14,457 (1995).
- [20]C. A. Ordonez, Phys. Plasmas, **1** 1359 (1994).
- [21]L. A. Schwager and C. K. Birdsall, Phys. Fluids B, **2** 1057 (1990).
- [22]C. A. Ordonez, Phys. Fluids B, **4** 778 (1992).
- [23]L. A. Schwager, Phys. Fluids B, **5** 631 (1993).

- [24]K. Kurihara, Y. Kiwamoto, T. Saito, K. Yatsu, and S. Miyoshi, J. Phys. Soc. Jpn., **61** 3153 (1992).
- [25]T. Saito, Y. Kiwamoto, K. Kurihara, T. Cho, M. Inutake, S. Miyoshi, T. Tamano, and K. Yatsu, Phys. Fluids B, **5** 866 (1993).
- [26]Y. Tatematsu, Y. Kiwamoto, T. Saito, and T. Tamano, J. Phys. Soc. Jpn., **63** 558 (1994).
- [27]Y. Tatematsu, Y. Kiwamoto, T. Saito, Y. Yoshimura, T. Takahashi, I. Katanuma, M. Inutake, and T. Tamano, J. Nucl. Mater., **220-222** 575 (1995).
- [28]T. Saito, Y. Kiwamoto, Y. Tatematsu, Y. Yoshimura, T. Takahashi, M. Inutake, and T. Tamano, Phys. Plasmas, **2** 352 (1995).
- [29]C. A. Ordonez and R. E. Peterkin, Jr., J. Appl. Phys., **79** 2270 (1996).
- [30]P. C. Stangeby in *Physics of Plasma-Wall Interactions in Controlled Fusion*, ed. D. E. Post and R. Behrisch (Plenum, New York, 1986) p. 41.
- [31]R. N. Franklin and W. E. Han, Plasma Phys. and Cont. Fus., **30** 771 (1988).
- [32]T. Intrator, M. H. Cho, E. Y. Wang, N. Hershkowitz, D. Diebold, and J. DeKock, J. Appl. Phys. **64**, 2927 (1988).

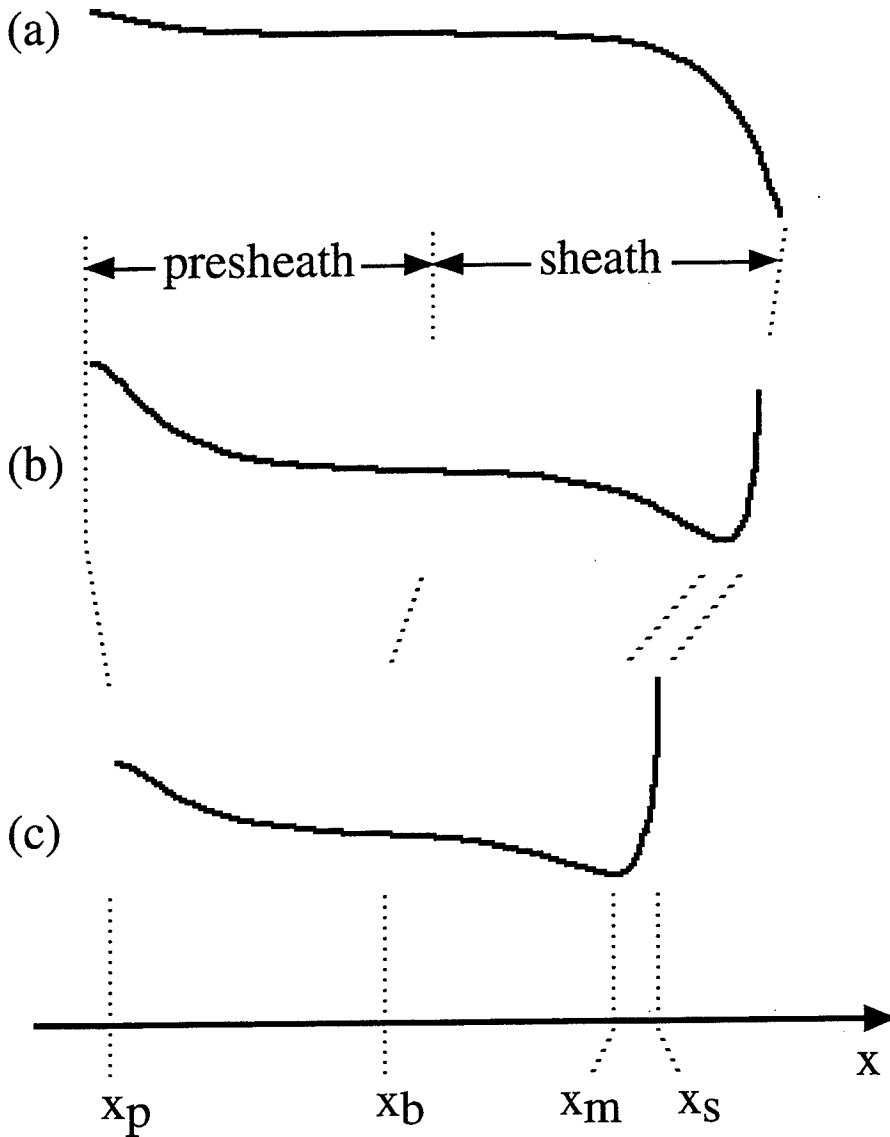


Fig. 1. Illustrations of possible electric potential profiles (solid curves) in the presheath and sheath for (a) no space-charge saturation, (b) space-charge saturation and a surface floating potential which is negative with respect to the edge plasma (at $x \leq x_p$), and (c) space-charge saturation and a positive surface floating potential. Although not apparent as a result of the limited resolution of the plots, the electric field is zero at x_p and x_b . The dotted lines are provided to guide the eye.

FURTHER ANALYSIS OF KILOHERTZ ORDER WAVES ASSOCIATED WITH
ELECTRON BEAM OPERATIONS ON STS 46

Michael J. Pangia
Associate Professor
Department of Geology & Physics

Georgia Southwestern State University
800 Wheatley Street
Americus, GA 31709-4693

Final Report for:
Summer Faculty Research Program
Phillips Laboratory / Hanscom AFB, MA

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

September 1996

FURTHER ANALYSIS OF KILOHERTZ ORDER WAVES ASSOCIATED WITH ELECTRON BEAM OPERATIONS ON STS 46

Michael J. Pangia
Associate Professor
Department of Geology & Physics
Georgia Southwestern State University

Abstract

This is a continuation of previous work to explain the 0.5 to 2 kHz frequency electron modulations observed by the Space Particle Correlator Experiment (SPACE) during beam operations on STS 46. It was determined that the ions flowed past the orbiter too quickly to be generated by a conventional beam instability. New information regarding the nature of the waves suggest that the wave could be a standing wave phenomenon, whose frequency is determined by boundary conditions imposed by the cross-sectional edge of the beam and an additional boundary. It is suggested that the additional boundary is the mach cone caused by the speed exceeding the ion sound speed. A 1-D problem is analyzed and is seen to not be a physical solution. It is conjectured that a standing wave solution might exist for the 2-D problem.

A suggestion for additional study is made in the report.

FURTHER ANALYSIS OF KILOHERTZ ORDER WAVES ASSOCIATED WITH ELECTRON BEAM OPERATIONS ON STS 46

Michael J. Pangia

I. Introduction

The previous work by *Pangia* [1995] considered a plasma instability as the mechanism for the strong electron modulations in the 0.5 to 2.0 kHz frequency range observed by the Space Particle Correlator Experiment (SPACE) during beam operations on STS 46. However, reconsideration of the basic premise shows that this is unlikely to explain the activity. With the orbiter speed being 7.8 km/s and an estimate of 3 m as the size of the instability region, the shuttle passes the region in about 0.4 ms. Based on the analysis by *Feng et al.* [1992] and the ion cyclotron frequency being 30 Hz at shuttle altitude, the maximum growth rate is 500 s^{-1} , which is insufficient to produce the observed wave amplitudes.

Making use of additional analysis of the data, the present work suggests that the waves might be standing waves caused by the effects of plasma boundaries that arise from the presence of the electron beam and perhaps the mach cone it produces.

II. Key Features of the Data

A current analysis of the data was done by *Gough et al.* [1996]. Here the key features are given:

- 1) The waves are present in situations where the magnetic field is nearly perpendicular to the ram velocity of the orbiter.
- 2) Although the frequency of the modulations can range from 0.5 to 2.5 kHz, they are discrete; specifically, they have a frequency spread of less than the lowest resolution of

SPACE, which is 39 Hz.

- 3) The frequencies tend to be higher for middle range beam pitch angle (60° - 120°), which, due to orientation of the detectors, corresponds to sampling particles off to the side of the beam column.
- 4) The waves are only observed downstream of the beam. In particular, *Feng et al.* [1992] showed that to a distance of about 400 meters downstream of the beam, the waves appear to be coming from the beam region and appear to be ion-acoustic waves.
- 5) The order of magnitude of the modulation is the inverse of the transit time of the orbiter traversing a region the size of the beam diameter.

III. Basic Premise

Producing a narrow band wave (Feature #2) is a challenge for most mechanisms. One possibility that comes to mind is that this is just a property that standing waves have. The question is how do standing waves fit into the existing situation. Firstly, it must be understood that the wave would have a standing pattern in the shuttle frame. Secondly, boundaries must be identified to contain the wave, thereby discretizing the frequency of the wave. One such boundary is the edge of the electron beam itself. Theoretically, another boundary that could exist is the mach cone produced by the supersonic motion of the shuttle (and hence the beam) through the plasma.

Figure 1 shows the basic layout of the boundaries and possible wave modes. The magnetic field would be perpendicular to the beam diameter. Note that the mach cone is detached from the beam by some stand-off distance, which is a necessary assumption for there to be two boundaries. Calculation of the actual shape of the mach cone and its stand-off distance would be

the subject of future research, as discussed in Section V.

Three wave patterns are drawn in Figure 1 with the regions labeled A, B and C. Waves in regions A and C would belong to the same standing wave mode, whereas the wave in region B may or may not, depending on what the eigenfunctions are for the geometry. However, the electron modulation

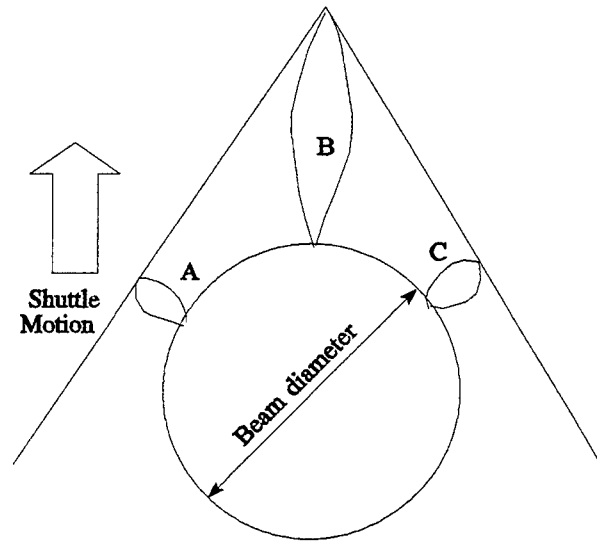


Fig. 1

that is observed by SPACE would depend on what region the electrons that SPREE is detecting are in. This allows for there being a dependence of modulation frequency on beam pitch angle (Feature #3), because, for mid-range beam pitch angles, SPACE would sample electrons in regions A or C, but not region B.

IV. One-Dimensional Analysis

Although Figure 1 is inherently two-dimensional, the geometry is too complicated. Instead a 1-dimensional simplification will be considered and is shown in Figure 2. The circles drawn with dashed lines at boundaries of the wave indicate that the boundary

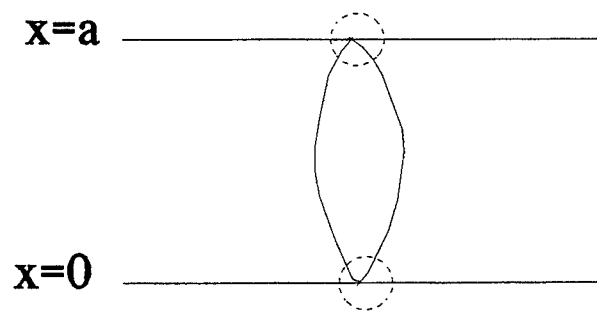


Fig. 2

conditions have to be applied there. The $x=0$ and $x=a$ lines represent the boundaries of the beam edge and the mach cone shown in Figure 1.

IV. A. Boundary Conditions

Common boundary conditions used in usual standing wave problems are zero amplitude or zero slope at the boundaries. However, neither would be true for a plasma due to Debye shielding that occurs, mostly by the electrons since only their thermal speed is significant compared to the orbiter speed [Krall and Trivelpiece, 1975]. In an electron plasma, the electric potential near a region of net charge will be

$$\Phi = \Phi_0 \exp(-|x|/\lambda_e) \quad (1)$$

where λ_e is the electron Debye length. This potential satisfies the following differential equation

$$\frac{d\Phi}{dx} \pm \frac{\Phi}{\lambda_e} = 0 \quad \text{for } \text{sign}(x) = \pm 1 \quad (2)$$

Therefore, the boundary conditions that will be used for Figure 2 will be based on (2). Specifically, they are

$$\frac{d\Phi}{dx} - \frac{\Phi}{\lambda_e} = 0 \quad \text{at } x=0 \quad (3)$$

$$\frac{d\Phi}{dx} + \frac{\Phi}{\lambda_e} = 0 \quad \text{at } x=a \quad (4)$$

IV. B. Equations

As the existence of standing waves does not depend on there being an instability, it is adequate to use the fluid equations for the ions and electrons. Recalling the basic premise that the waves are standing waves in the shuttle frame, the equations will be written in that frame.

Various approximations can be made that simplify the problem considerably. Firstly, the modulation frequency is much less than the electron plasma and gyro- frequencies, so inertial effects and gyro-motion of the electrons can be neglected. The ion gyro-frequency is much less than the modulation frequency, so the gyro-motion of the ions can also be neglected. The temperature of the ions can also be neglected in comparison to the electron temperature. In the shuttle frame, the ions will have a flow speed of V_s , the shuttle speed, and their direction of motion will be towards decreasing x . With these simplifications, the equations are:

$$0 = en_o \frac{\partial \Phi}{\partial x} - T_e \frac{\partial n_e}{\partial x} \quad (5)$$

$$m_i n_o \left(\frac{\partial v_i}{\partial t} - V_s \frac{\partial v_i}{\partial x} \right) = -en_o \frac{\partial \Phi}{\partial x} \quad (6)$$

$$\frac{\partial n_i}{\partial t} - V_s \frac{\partial n_i}{\partial x} + n_o \frac{\partial v_i}{\partial x} = 0 \quad (7)$$

$$-\frac{\partial^2 \phi}{\partial x^2} = 4\pi e(n_i - n_e) \quad (8)$$

where the equations are the linearized versions of the electron and ion momentum equations, the ion continuity equation and Poisson's equation, respectively. The electron and ion number density was linearized about n_o with n_e and n_i being the first-order deviations from equilibrium for the respective species. The temperature T_e is that of the electrons. The velocity v_i is the linearized ion flow velocity. Eqs (5) - (8) can be combined to obtain

$$-\left(\frac{\partial}{\partial t} - V_s \frac{\partial}{\partial x}\right)^2 \frac{\partial^2 \Phi}{\partial x^2} + \left(\frac{\partial}{\partial t} - V_s \frac{\partial}{\partial x}\right)^2 \frac{\Phi}{\lambda_e^2} - \omega_i^2 \frac{\partial^2 \Phi}{\partial x^2} = 0 \quad (9)$$

where ω_i is the ion plasma frequency.

IV. C. Standing Wave Analysis

Standing wave solutions in 1-dimension would have the form

$$\Phi(x,t) = Af(t) \cdot \phi(x) \quad (10)$$

where the f and ϕ are sinusoidal functions and A is a constant. A general way to write $f(t)$ and $\phi(x)$ is

$$f(t) = \cos(\omega t) + B_1 \sin(\omega t) \quad (11)$$

$$\phi(x) = \cos(kx) + B_2 \sin(kx) \quad (12)$$

where B_1 , B_2 , ω , and k are additional constants, which would have to be related by the boundary conditions and (9).

Application of boundary condition (3) gives that $B_2 = 1/(k\lambda_e)$. Application of boundary condition at $x=a$ would be difficult to solve, but an approximation concerning wave number, k , can be made. Namely, the wavelength of the wave, as seen from Figure 2, is on the order of a , which in turn is on the order of the electron gyro-radius, which is considerably larger than the electron Debye length; therefore $k\lambda_e \ll 1$. In this approximation, (4) gives

$$k \approx n \frac{\pi}{a}, \quad \text{for } n=1,2,3,\dots$$

Substitution of (10) into (9), with B_2 and k already determined, would in principle give the remaining two constants, ω and B_1 . However, this method fails by yielding an imaginary value for B_1 , which is unphysical.

V. Discussion

The failure of a 1-D standing wave solution existing is possibly due to the problem being inherently two-dimensional. If standing wave solutions did exist, all of the key features of the data would be explained. As mentioned in Section II, the discreteness of standing wave frequencies would account for Feature 2. Furthermore, it is expected that the frequency would be of the order of V/a . This would agree with Feature (5). Feature 3 could be accounted for by the fact that, in Figure 1, regions A and C have a smaller distance between the boundaries and therefore a higher frequency. The feature observed by *Feng et al.* [1992] (Feature 4) is explained by recognizing that once the ions leave the region they flow downstream and are free to oscillate at their natural frequency.

VI. Future Steps

Future research is in order, as no definite answer has yet been found. One prospect is to solve the standing wave problem in two-dimensions. However, there is a major question as to the shape and location of the mach cone. Therefore, it is felt that the best step to take next is to solve the problem by computer using a particle simulation of the region. Assuming the simulation is capable of describing the necessary physics, the development of the mach cone as well of the wave effects can be studied.

References

- Feng, W., D. A. Gurnett, and I. H. Cairns, Interference patterns in Spacelab 2 Plasma Wave Data: oblique electrostatic waves generated by the electron beam, *J. Geophys. Res.*, 97, 17,005 1992.
- Krall, N. A., and A. W. Trivelpiece, *Principles of Plasma Physics*, McGraw-Hill, New York, 1975.
- Pangia, M. J. Set-up for plasma instability analysis of kilohertz order waves associated with electron beam operations on STS 46, *1995 AFOSR Summer Research Final Reports*, p. 31-1, 1995.
- Gough, M. P., D. A. Hardy, W. J. Burke, M. R. Oberhardt, L. C. Gentile, C. Y. Huang, B. McNeil, K. Bounar, D. C. Thompson, and W. J. Raitt, Heating and low-frequency modulation of electrons observed during electron beam operations on TSS 1, *submitted to J. Geophys. Res.*, 1996.

TEMPORAL-DISPLACEMENT STEREOGRAMS OF THE IONOSPHERE:
AN EXPLORATION OF THEIR UTILITY IN THE ANALYSIS OF EQUATORIAL
EMISSION DEPLETION BANDS.

Ronald M. Pickett
Professor
Department of Psychology

University of Massachusetts Lowell
One University Avenue
Lowell, MA 01854

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

September 1996

TEMPORAL-DISPLACEMENT STEREOGRAMS OF THE IONOSPHERE:
AN EXPLORATION OF THEIR UTILITY IN THE ANALYSIS OF EQUATORIAL
EMISSION DEPLETION BANDS.

Ronald M. Pickett
Professor
Department of Psychology

Abstract

Optical images of the night sky are a primary source of information in research on the ionosphere. Typically only one photometer is employed, and the resulting single-point-of-view images are analyzed primarily for information on the transverse size, shape and position of the structures of interest. The present project is concerned with exploring the potential of getting volumetric information from the optical images via optical stereograms. To obtain true stereograms would require employing two photometers an appropriate distance apart on the ground, each aimed at the same section of sky and each recording an image at the same time. We sought to approximate that ideal situation with images obtained by just a single photometer. We created temporal-displacement stereograms and examined them for the light they might throw on the potential utility of true stereograms. We concluded that they do not look very promising as a source of volumetric information in their own right, and they are not a sufficient basis for evaluating the potential informativeness of true stereograms.

TEMPORAL-DISPLACEMENT STEREOGRAMS OF THE IONOSPHERE: AN EXPLORATION OF THEIR UTILITY IN THE ANALYSIS OF EQUATORIAL EMISSION DEPLETION BANDS.

Ronald M. Pickett

Introduction

Optical images of the night sky are a primary source of information in research on the ionosphere. Ground-based all-sky photometers record changes in the shape and position of structures of interest, which can then be related to other structures and events elsewhere in the ionosphere and larger earth environment. Typically these images are obtained by a single photometer, and the resulting single-point-of-view images are analyzed primarily for information on the transverse size, shape and position of the structures of interest. Information on volumetric characteristics of the structures is obtained by other means, like radar sounding, and typically only very sparse volumetric data are obtained. The present project is concerned with exploring the potential of getting volumetric information from the optical images via stereograms. For these exploratory studies, we focus on a particular type of ionosphere structure -- equatorial emission depletion bands. Depending on the season and locus of observation, the ionosphere will be found to visibly glow over broad areas of the night sky. Near the magnetic equator, various chemical processes conspire to reduce this air glow in bands that run north to south and move smoothly west to east over the night sky. These equatorial emission depletion bands are of great interest as a source of information on which to build models of the ionosphere. They are also a subject of intense inquiry because they can severely disrupt radio communications, including line of sight signals between the ground and navigational satellites. Thus, it is important to understand as much as possible about how these structures grow, move and penetrate up into the ionosphere.

As explained in more detail below, to obtain true stereograms would require employing two photometers an appropriate distance apart on the ground, each aimed at the same section of sky and each recording an image at the same time. Those two images, taken from just the right pair of

angles would comprise a visually fusible stereogram, a potentially valuable source of volumetric information. Our challenge here is to somehow approximate that ideal situation with images obtained by just a single photometer. To do that we take advantage of the situation, particularly convenient with equatorial depletions bands, that they translate smoothly west to east over the photometer. We can, thus, obtain two images of the same banded region from two different angles, but instead of obtaining two images, each taken from a different position at the same time, we obtain two images each taken from the same position at two different times. If the depletions did not change in size, shape or altitudinal position as they traversed across the sky, this temporal-displacement technique would be exactly equivalent to that achieved with two spatially separated photometers. The approximation enters because the depletions do change as they translate. The big question is whether they change little enough to provide an informative approximation to the ideal.

This project was aimed at creating temporal-displacement stereograms and examining them for the light they might throw on the potential utility of true stereograms for viewing and measuring volumetric properties of emission depletion bands. In this report we first present technical background on the all-sky images and then on the technique of creating stereograms.

Technical Background

All-sky Photometer System and Images. The all-sky photometer records the images through a fish-eye lens onto a 512 x 512 CCD array. It takes a series of four images per minute, one unfiltered and one through each of three narrow-band filters -- at 557.7nm, 630.0nm and 777.4nm. For our examples, we use just the 630.0nm images, but the system we have built for creating and displaying the stereograms was designed to handle all three narrow-band images both individually and in color integration.

The raw images are distorted by the fish-eye lens. The area of sky covered by one pixel varies greatly across the image, growing progressively larger from zenith to the horizon. The images can

be transformed to correct for that optical distortion but only with some significant loss in quality at the edges where the required stretching results in strong, artifactual striations. In our examples we present stereograms just of the transformed (linearized) images. Moreover, to avoid the effect of the striations at the periphery, which we think interfere with the perception of depth, we have cropped away the edges, leaving just the central 256 x 256 pixels of the transformed image.

Temporal-Displacement Stereograms. A stereogram of an object is usually created by taking two pictures of it from the positions at which the two eyes would view it in direct observation. The two pictures are then displayed side by side, and an optical device is provided to assist the observer in fusing the two images. If the two images are taken at positions farther apart than the distance between the two eyes, up to a point the observer can still effectively fuse them into a single image, but the effect of increasing the inter-eye distance is to increase the apparent depth of the object, making it look deeper than it truly is. Making the distance less than the inter-eye distance results in the opposite distortion, making the depth of the object appear to be less than it truly is.

With an effect equivalent to moving the camera from one eye position another, one can hold the camera position constant and move the object laterally. The object is moved to the left a distance sufficient to acquire an image of the object as it would be seen by the right eye and to the right to acquire an image as it would be seen by the left eye. It is by that logic, that we acquire the image pairs for time-displacement stereograms. If we consider the time at which the banded region of interest is directly overhead as $t=0$ minutes, then the left eye image for the stereo pair would be the one recorded at $t + m$ and the right eye image would be the one recorded at $t - m$. As m is increased, the apparent depth of the image volume is increased up to the limit at which the disparate images can be effectively and comfortably fused into one image.

Our measure of the temporal displacement of the image pair, $\pm m$, is related to the effective spatial displacement by the estimated speed of translation of the banded region of interest over the photometer. We estimate the speed at which depletions move across the sky as approximately 4.8 km per minute. So, for each minute back in time and forward in time, the zenith scene of interest is

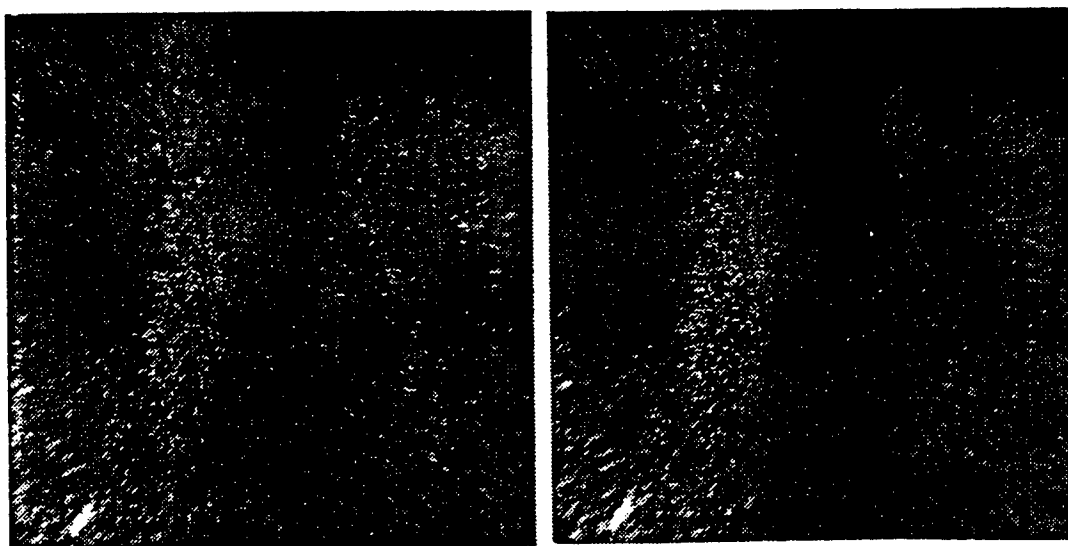
effectively slid to the left/right 4.8 km. That would be the same distance required for two photometers on the ground to be moved to the left/right to acquire an equivalent stereogram of the zenith scene at $t=0$.

Results

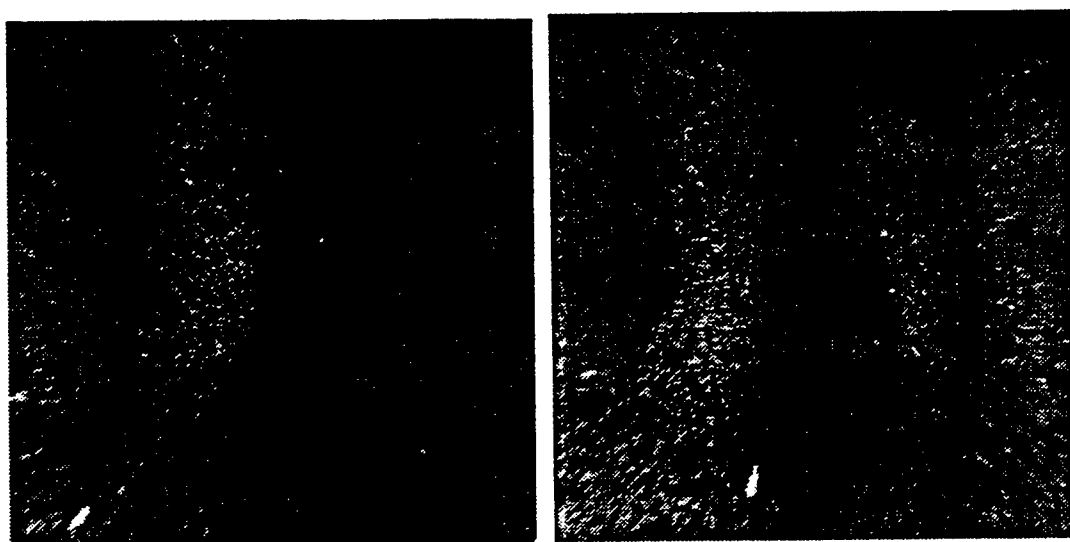
In the course of this project, we created and examined numerous stereograms of depletions in available records. We show in Figure 1 a typical case, two stereograms of a zenith scene of two emission depletion bands, one band about in the middle of the image, the other entering from the left. For figure 1a, the two images are identical ($t=0$) and the stereogram provides a baseline display of zero depth perception of the scene. For Figure 1b, $t = \pm 8$. It is evident that the scene in the left eye image is translated to the right and that in the right eye to the left, providing a potentially strong basis for perceiving the scene in depth. Given that each minute of temporal displacement corresponds to 4.8 km of spatial displacement, this stereogram is effectively equivalent to one taken with two photometers 76.8 km apart.

Fusion of stereograms, even of sharply contoured, high contrast scenes is not easy for many people. An optical aid should be employed for starters, but some persistence and patience is required. The $t = 0$ stereogram is easy to fuze, because even though it is very ambiguous and noisy, the noise elements are identical between the two images and, thus, provide a strong basis for fusion. But the $t = \pm 8$ stereogram is a challenge to fuze even for those who have strong stereoscopic fusion capability.

When well fused, the $t = \pm 8$ stereogram presents at best a very ambiguous picture of the depletions in depth. A weak impression of depth is achieved, with the bright bands of air glow coming out and the depletion bands receding. But this impression is quite weak, and very little basis is provided for gauging how informative true stereograms would be. One encouraging suggestion comes from the fact that just because the scene is amorphous and noisy may not preclude achieving strong fusion. Noisy variations in the scene, provided they are intrinsic to the



(a)



(b)

Figure 1. Temporal-displacement stereograms. Shown in (a) is a stereogram with $t = 0$ minutes of displacement. The left and right images are identical, and it serves as a baseline display with zero depth. Shown in (b) is a stereogram of the same zenith scene with $t = \pm 8$ minutes of displacement. The stereogram in (b) is effectively equivalent to a true stereogram obtained with the photometers 76.8 km apart.

scene and not added by the imaging system, might well provide an effective basis for fusion of images taken from different positions at the same instant, since they would be highly correlated.

Conclusions

While we have conducted only a very limited exploration of what temporal displacement stereograms can show, we can conclude that they do not look very promising as a source of volumetric information in their own right. If the general idea of using stereoscopic displays to obtain volumetric information looks interesting, it would probably be more appropriate to collect some true stereograms and study them rather than pursue this temporal-displacement technique any further.

LOW VOLTAGE ANALOG CIRCUIT DESIGN FOR RADIATION TOLERANCE

Edgar Sánchez-Sinencio
Professor
Department of Electrical Engineering

Texas A&M University
College Station, TX 77843-3128

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC
and
Phillips Laboratory

August 1996

LOW VOLTAGE ANALOG CIRCUIT DESIGN FOR RADIATION TOLERANCE

Edgar Sánchez-Sinencio

Professor

Department of Electrical Engineering

Texas A&M University

Abstract

A low voltage analog circuit design approach for radiation tolerance is investigated. Analog circuits are sensitive to process variations, temperature, low voltage power supply and the total dose radiation. Nevertheless analog circuits can not be avoided in any modern integrated system. Even a conventional digital signal processing (DSP) system requires a preamplifier and an A/D converter and often also a D/A converter. Mixed-mode circuits involving digital and analog circuits are standard for modern circuit design. In this report a method for radiation effect MOS transistors modeling using conventional circuit simulators such as PSPICE is introduced. Furthermore an adaptive scheme to correct the circuit performance deviations due to the radiation effects is presented. Preliminary simulated results confirm the correct radiation effects modeling as well as the potential application of adaptive schemes to correct performance perturbations due to radiation.

LOW VOLTAGE ANALOG CIRCUIT DESIGN FOR RADIATION TOLERANCE

Edgar Sánchez-Sinencio

I. Introduction

The design of modern electronic systems for space and military applications demands consideration of the following effects on system performance:

- I) Technological process variations.
- ii) Temperature range.
- iii) Low voltage power supply.
- iv) Transient and total dose radiation.

Most of the practical integrated circuits found in applications involve both digital and analog circuits. These circuits are often referred as mixed-signal. In the last 20 years a number of important results have been reported on hardened digital circuits. Schemes for high radiation tolerance using dedicated technology, clever layouts and circuits techniques have yield fairly reliable digital circuits and systems. Much less attention has been paid to the design of analog circuits for high radiation tolerance. In addition as technologies scale down, the supply voltage must scale down. The typical 5 V for standard 1 μm technologies is reduced to about 3 V for 0.5 μm , and even smaller for 0.25 μm . As long as analog and digital circuits are merged on common VLSI silicon substrates designers must learn to design analog functions with low-voltage power supply. From the point of view of power consumption minimization there is no motivation to reduce the supply voltage of analog circuits. In fact, it can be shown that in order to preserve the same dynamic range (DR) while reducing the supply voltage, the power consumption must be increased. This is because to keep the same DR the capacitors of the system should be scaled up with the square of the supply voltage scaling factor.

In this report we focus on the design of analog circuits for radiation tolerance. The approached is constrained to:

- a) A conventional circuit simulator such as PSPICE.
- b) Use of a standard CMOS process technology such as the ones available from the MOSIS (MOS Implementation Service) clearing house developed by DARPA and managed by ISI (Information Science Institute).

- c) Power Supplies equal or less than 3.3 volts.
- d) Take into consideration radiation effects and implement a circuit scheme to correct the performance deviations due to the radiation effects.

A space environment can affect CMOS devices by several primary radiation components, i.e., galactic cosmic rays, electrons, protons and solar flares. The primary effects in devices are ionization effects and Single-event effects. This environment imposes very difficult circuit and system design constraints. We will address only a few of the problems involved in designing microelectronics circuits in a space environment. A number of authors have concentrated in radiation tolerant digital circuits [1]- [4], [8]. In particular we will focus on low voltage analog circuit design since not enough results have been reported, in the literature, [9]-[12].

The outline of this report is as follows:

Section II describes how to incorporate radiation effects into PSPICE, and includes examples verifying the radiation effects on MOSFET modeling. Section III introduces the key issues on Low Voltage (power Supply) Analog Circuit Design. Several low voltage analog circuit building blocks are revised. Section IV is devoted to an Adaptive Algorithm. This algorithm is derived and used to correct radiation circuit deviation performance. In Section V, an example of a circuit which performance is perturbed by radiation and corrected by an adaptive scheme is illustrated. Section VI gives conclusions and outlines future work.

II. Radiation Effects Modeling on Diodes and MOSFET using PSPICE Simulator.

In order to evaluate the radiation effects on circuit devices we need to have a good simulator capable of identifying critical circuit components. There are a few simulators developed by industry which are capable of simulating radiation effects on circuits, unfortunately these software packages are either difficult to get or are prohibitively expensive for the academic environment. One public domain software package BERT [3]

that can help to design circuit under radiation effects is available from U. of California at Berkeley. Even though BERT was developed as a circuit-reliability simulator that predicts hot-carrier degradation, oxide reliability, and electromigration failures for integrated circuits undergoing dynamic operation can be used as a radiation reliability simulator. The emphasis of this simulator is for digital circuits.

Next, we will describe how to incorporate the radiation effects in diodes and MOSFETs in a conventional circuit simulator such as PSPICE.

II.1- Photocurrent generators associated with *p-n junction* diodes. The magnitudes of the photocurrents are frequently obtained from experimental data. If the fabrication process data is available, then theoretical expressions can be also obtained. The photocurrent expression due to the ionizing radiation [2,5] usually involves the integral of an *erf(.)* function . Fortunately this mathematical expression (and the experimental data) can be approximated to a pulse with exponential rise and fall (decay) times, thus allowing this mathematical description to be easily incorporated into a PSPICE input file. For the circuit analysis a photocurrent generator should be placed in parallel with every single diode of the circuit. These diodes can be parasitic or design component diodes. Fig. 1 illustrates the connection of the diode and the current generator. Note the direction of the current

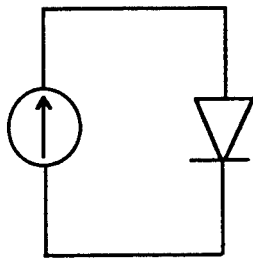


Fig. 1 Photocurrent Generator Placement

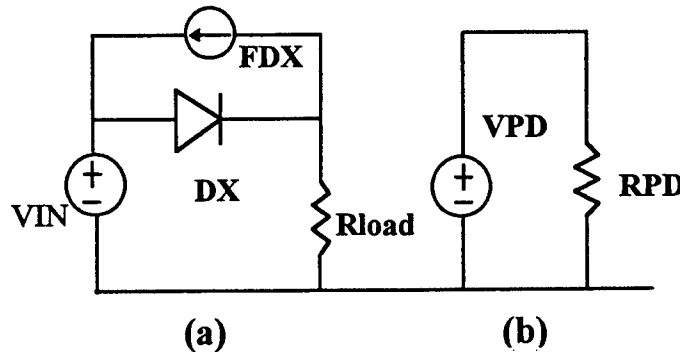


Fig. 2 Diode Photocurrent Generator Example (a) Circuit; (b) Auxiliary Generator

source is the same as the leakage current the diode. This approach is valid for any BJT, or any other transistor type that contains *p-n* junctions. To illustrate how the current generator is described in a PSPICE input file, we are going to use an example containing an input sinusoidal voltage source in series with a rectifying diode and a resistive load. The diode is connected in parallel with the photocurrent generator as in Fig. 1. The circuit is shown in Fig. 2 (a). The circuit in Fig. 2 (b) is used to generate the photocurrent implementation. The corresponding input PSPICE file is given next.

TITLE: DIODE HALF RECTIFIER WITH PHOTOCURRENT DUE TO RADIATION

* First the photocurrent generator to be placed across diode are generated

.PARAM DR = 1E7 ; initialization

.PARAM JPM = 4E-02 ; peak measured current density

.PARAM DM = 1E7 ; at measured dose rate

.FUNC JP(DR) { DR*JPM/DM} ; linear approximation of photocurrent

* Perform simulations at 1E6 ,1E7, 1.5E7 and 1E8 rad(Si)/sec

.STEP PARAM DR LIST 1E6 1E7 1.5E7 1.0E08

VPDI 99 0 EXP(0 {JP(DR)} 100NS 8NS 150NS 50NS)

RPDI 99 0 1.0

DX 1 88 DMODX ; Area=2.5E-03

* A current controlled current source, this photocurrent source is controlled by VPDI

FDX 1 88 VPDI -2.5E-03

RLOAD 88 0 10.0

.MODEL DMODX D (IS=1E-11)

VIN 1 0 SIN (0 900M 4E06)

.tran 1NS 500NS

.probe

.END

Note that VPDI is an exponential pulse with a variable peak voltage function of the dose rate with rise and fall time of 8 nanoseconds and 50 nanoseconds, respectively. VIN is a sinusoidal generator with a peak amplitude of 900 mV and a frequency of 4 MHz. The results of this PSPICE run are shown in Fig. 3. In Fig. 3(a) the half-wave rectifier shapes for different dose as well as the input voltage (scaled down) are shown. To have a better appreciation during the interval at which the diode ideally is off, in Fig. 3 (b) we zoomed between 101 nsec and 240 nsec. It can be noticed that for a low dose the behavior is very close to the ideal case, as the dose increases the deviations increase. In Fig. 3 (c) the

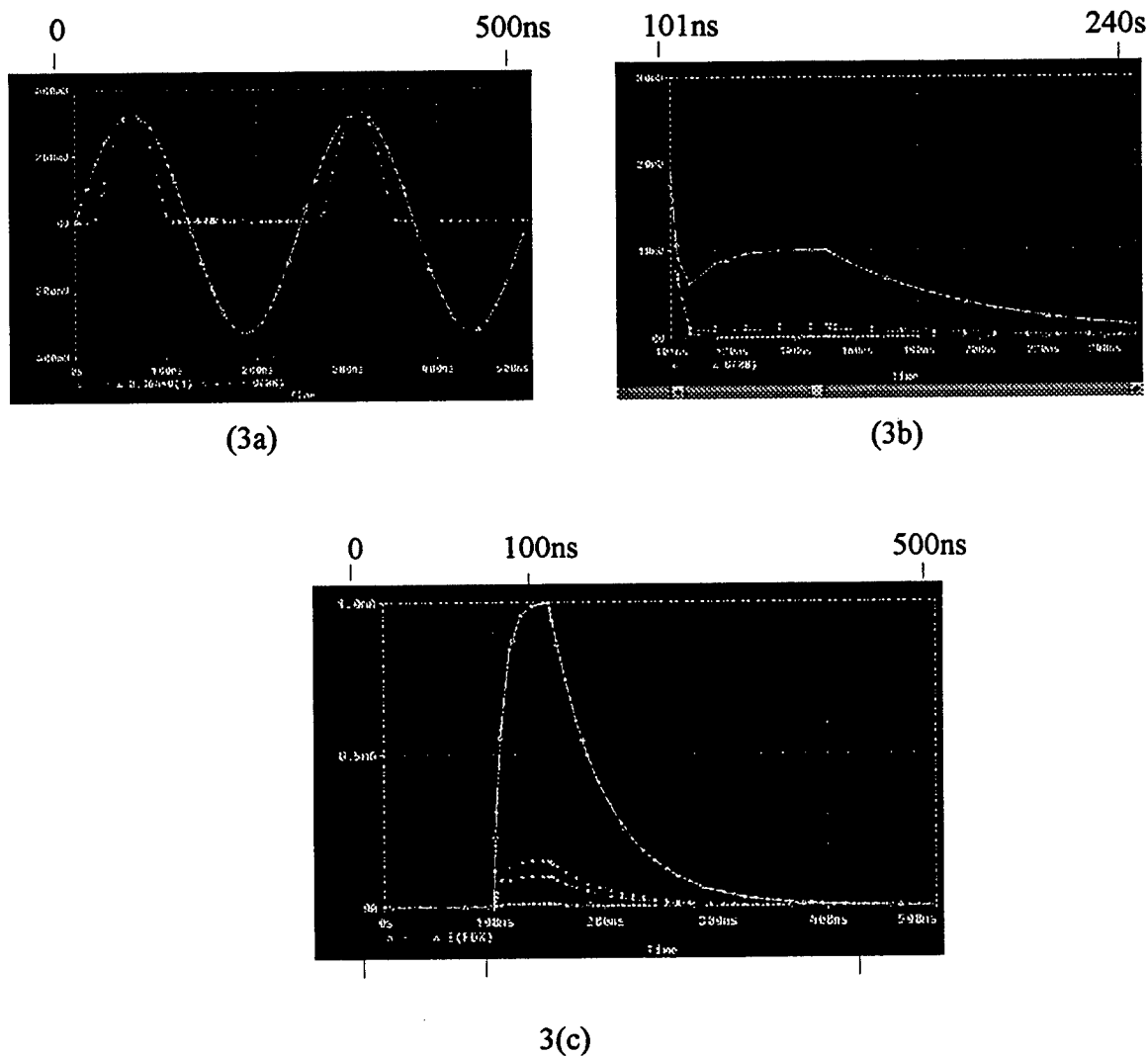


Fig. 3 SPICE Results of Circuit of Fig. 2.

photocurrent generator FDX shows the different curves for different dose. The modeling of photocurrent previously discussed is mainly a transient phenomena. To model leakage currents in MOS devices a similar approach could be followed. That is to add a current source across every p-n junction of the device. This static current source could be modeled as a function of the dose.

II.2 Total Dose Effects Modeling associated with PMOS and NMOS Transistors.

When high-energy protons and electrons pass through MOS structures, they produce electron-hole pairs within the gate and field oxides of the structures. The electrons resulting from ionization have high mobility in the oxide and are quickly swept out by internal fields. The holes have much lower mobility, and some fraction of these holes will be transported to the silicon/silicon-dioxide interface, where they will be trapped. This effect will change the threshold voltage and the mobility of the gate and the field-oxide transistors, varying their characteristics. Next we illustrate for circuit simulations [6]-[7] how to take into account the total dose effects, assuming the threshold voltage and mobility are the dominant factors for the transistor models. The threshold is characterized by a second-order polynomial (quadratic) approximation, thus the threshold voltage of a NMOSFET can have positive and negative values depending on the dose. The PMOS transistors in comparison with the NMOS transistors have threshold voltages that change slower with radiation. For n-channel devices the threshold shift is negative, causing a transistor to move toward depletion mode as the total dose increases. However, the threshold voltage values of p-channel structures are always negative. The reduction of mobility is the result of increased diffusion carrier scattering at the silicon surface due to interface states. Thus for circuit simulation the mobility is characterized by a straight line with negative slope. Next, an example for modeling the total dose effects in a NMOS transistor in a partial PSPICE input file is shown. Note that first, the coefficients values of the quadratic and linear functions are given, and then the functions are defined.

```
.PARAM DOSE = 0.0          ; initialization of total dose
.PARAM VTOI = 0.863        ; initial threshold voltage ( no radiation)
.PARAM UOI = 562.8          ; initial mobility ( no radiation)
.PARAM A1 = 4.18E-16        ; linear coefficient of the Vt function
.PARAM A2 = 2.76E-12        ; quadratic coefficient of the Vt function
* Perform simulations at 0, 1E5 and 1E6 rad(Si)/sec
.STEP PARM DOSE LIST 0 1E5 1E6
* Next the quadratic and linear functions of the threshold and mobility functions,
* respectively, are described.
```

```

.FUNC VTO_DOSE(DOSE) {VTOI*(1-A1*DOSE + A2*DOSE*DOSE)}
.FUNC UO_DOSE(DOSE) { UOI*(1-A3*DOSE)}
.MODEL NRAD NMOS (VTO={VTO_DOSE(DOSE)} UO={UO_DOSE(DOSE)})

```

Fig. 4 shows an example of a transconductance amplifier (an inverter using digital circuit terminology). In this plot it can be observed how the transfer characteristics move into depletion region as it was anticipated. Other examples involving MOSFET circuits under radiation will be given in section V, together with an adaptive scheme capable to correct the perturbations due to the radiation effects.

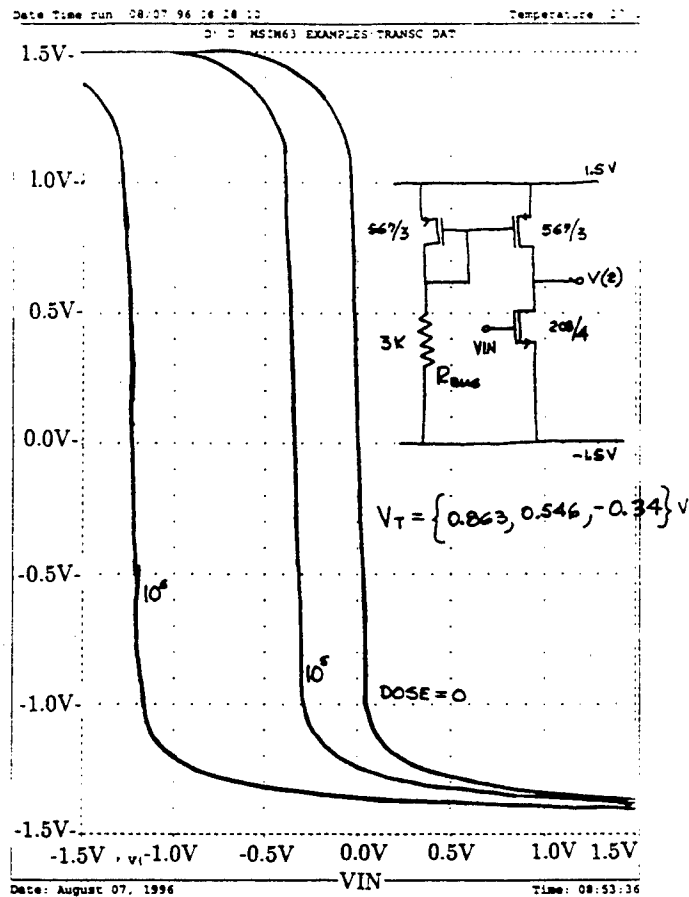


Fig. 4 Radiation Effects on Transconductor

III. Fundamental issues on Low Voltage Analog Circuit Design.

There is a clear tendency in IC design and fabrication, to employ lower voltage supplies and to push technologies to scale down. The most powerful circuits in practice are the so called mixed-mode signals circuits, these circuits consist of an optimal combination of digital and analog circuits merged on common silicon substrates. Thus, VLSI designers must learn to design analog function [14], [16] with low-voltage supply.

The modern design of analog circuits using low-voltage supply and/or low power consumption demand proper system architectures, suitable circuit models, trade-off optimization, and innovative circuit design. The critical issues when designing analog circuits with low-voltage supply are diverse. One of the first problems to tackle is the fact that to obtain, say, large voltage gains in an amplifier, large output impedance are required.

Fig. 5 shows different types of architectures for a simple inverting amplifier. The bias current sources are assumed to be ideal, that is $R_o = \infty$. The voltage gains for the circuits are $-g_m R_i$, $i = a, b, c$. Where g_m is the transconductance gain of the driver transistor M1, and R_i is the output impedance of the different configurations. Fig. 5(a) is a one (driver) transistor with an ideal bias current source, its simplest transistor implementation is shown in (b). The voltage gain is about $-g_m/(g_{o1} + g_{o2})$. In a typical process technology this yields gain between -30 and -60 or

less. The minimum voltage supply is about $\min \{V_{DSAT1} + |V_{DSAT2} V_{DSAT1} + V_{Tn}\}$, where V_{DSAT} is the minimum ($V_{gs} - V_t$) to keep the transistor operating under saturation, and V_{Tn} is the N-MOSFET threshold voltage. This circuit realization is suitable for low-voltage supply, unfortunately the voltage gain is often not sufficient for many practical applications. Thus other alternatives need to be searched. Fig. 5(c) is a circuit with large output impedance, and the gain can be increased by a factor of 10 to 100 with respect to circuit of Fig. 5(b), however, the minimum voltage-supply has to significantly increase higher than about $4V_{DSAT}$. Fig. 5(d) is a folded cascode circuit which can operate with lower supply voltage while maintaining a relative large (output impedance) voltage.

However, these simple amplifiers need to be modified to handle common mode signals. This requirement implies the use of differential amplifiers. It is known that for low-supply voltage a differential pair has limited input common-mode voltage. Thus, other alternatives need to be explored to obtain rail-to-rail common-mode voltage swing.

Recent trends in analog circuit design for low-supply voltages involve:

- i. Use of parallel NMOS and PMOS differential pairs [14] to handle large common-mode input signals.
- ii. Use of floating gate techniques [14] by decreasing (internally) the input signal and thus improving linearity.
- iii. Use of current-mode techniques [16] to have better control of large process variations.
- iv. Use of pseudo-differential mode circuit which does not involve the source bias current of differential pairs. This allows it to handle low-supply voltages.
- v. Use of simple cascade amplifiers in multiple feedback (nest) configurations.

The design constraints of analog circuit for radiation tolerance can be reformulated as analog circuit design for large process variations. Of particular importance are the variations of threshold voltages, mobility and offset current. This is of critical importance when the process technology has not been developed for hardened electronics.

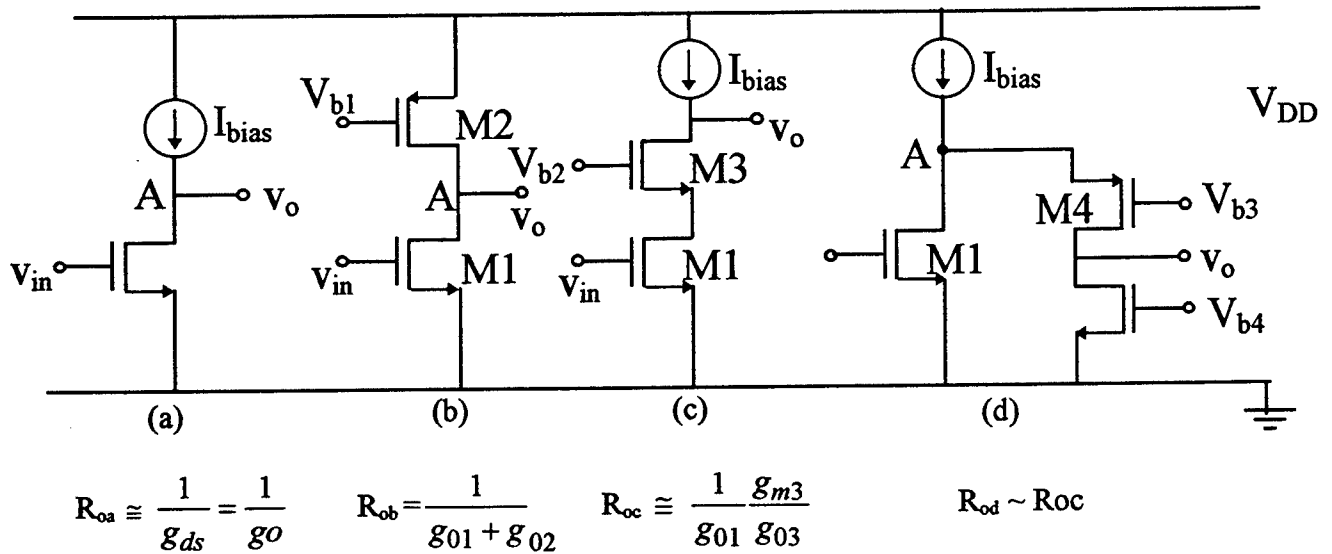


Fig. 5 Simple Amplifiers

IV. A Correction Scheme for Circuits Perturbed by Radiation.

Analog circuit performance is mainly sensitive to technological process variations, temperature, layouts, component matching and low voltage power supply. Furthermore when the analog circuits are exposed to radiation their performance are perturbed. The perturbation will depend on the amount of total dose radiation and annealing effects. Several designers have tackled this problem by using a specialized layout [1], and/or by using radiation hardened process technology. Also for digital circuits, designers can change the aspect ratios between p-channel and n-channel devices. For example in a typical CMOS inverter the ratio between (W/L)s of the NMOS and PMOS transistors is given by K_{Pn}/K_{Pp} , where $K_P = \mu C_{ox}$. Now by changing this ratio which is typically 2 to a higher value, say 3 or 4, the tolerance to threshold voltage can be increased. Radiation tolerance is also dependent on the circuit architecture, it is known that digital designers prefer for three or more inputs the CMOS NAND logic gate instead of a CMOS NOR gate for hardened circuit purposes. This preference is due to the better sensitivity to leakage currents of the NAND gate, i.e., for the NOR gate, the increased channel-leakage currents through the parallel connected NMOS transistors will increase static power dissipation when all inputs are "low".

The design of analog circuit can follow similar layouts and architectural schemes as well as using hardened process technology to reduce the radiation effects. However, the design of analog circuits is more complex than designing digital circuits, since threshold perturbations result in different bias conditions and this consequently changes the small signal parameters (which determine gains, impedances, etc). If the specialized hardened process technology is not available, then the circuit design complexity increases even more.

In this section an approach to compensate the analog circuit perturbed by radiation effects is presented. The perturbations can be not only due to radiation but also to any other type such as temperature, offset voltage (or current), process variations and component mismatching. The proposed scheme is based on an adaptive Least-Mean-Squares (LSM) algorithm. This algorithm minimizes the mean-square error between the desired response

and the actual response of the system (or circuit). It minimizes the error by updating the system coefficients through a feedback loop. Fig 6 shows the block diagram of the LMS algorithm. For the sake of completeness the LMS Algorithm derivation is included in an appendix at the end of the report. The algorithm consists on the implementation of the first-order differential \dot{e}

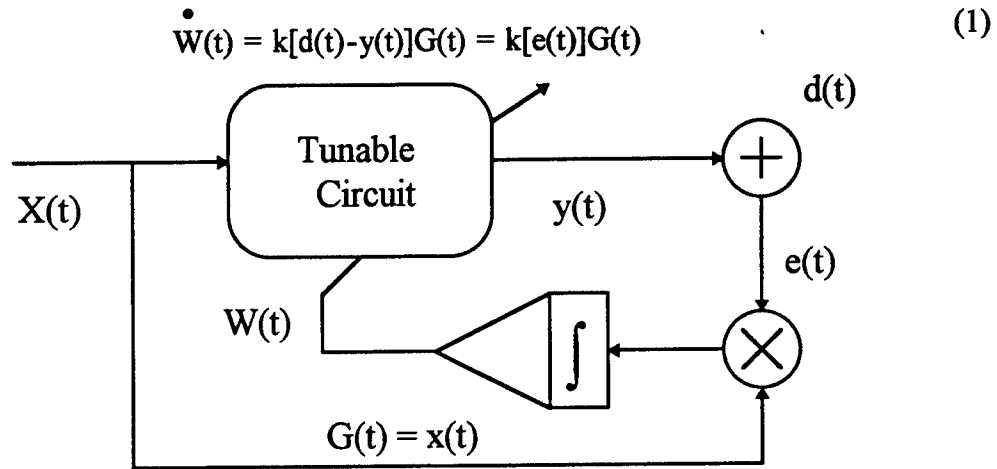


Fig. 6 A Block Diagram Implementation of the LMS Algorithm

Where $W(t)$ is the correcting (tuning) signal, k is an adaptation constant, $d(t)$ is the desired circuit output, $y(t)$ is the actual circuit output, $G(t)$ is the tuning gradient (partial derivative of $y(t)$ with respect to $W(t)$). In the particular case of a linear system $G(t) = x(t)$ as shown in the appendix. $X(t)$ is the input signal. The desired output depends on the application. For instance in the example to be shown in section V, the desired output is $A_v \cdot V_{in}$, where A_v is the voltage gain at the frequency of interest, and V_{in} is the sinusoidal input signal with the frequency of interest. Ideally, there is no restriction on the desired output which can be a phase delay, a peak voltage, a group delay, etc. The interesting issue is the implementation of the device to measure (sense) the actual output, $y(t)$. For high frequency the device to measure the actual output can be difficult to implement, alternatives to make easier the implementation and the measurement rely on AC to DC converters, peak detectors, or digital phase detectors. Another equally

important issue is the implementation of the desired output, this sometimes can take the form of a DC signal or by using the signal of the digital clock which typically is very accurate since is based on crystal oscillators. This signal from the clock can be an excellent stable reference.

Another factor very important in practical IC is the accurate implementation of the first-order differential equation discussed before. This is still an open problem in the literature. For stringent applications, this author suggests, as one of the possible solutions, a digital implementation of this equation, which can be converted into a difference equation. This will be further discussed in the section of conclusions and future work.

V. A Radiation Transconductance Amplifier Compensated by an Adaptive Scheme.

Fig. 7 shows a transconductance amplifier which is the complementary of circuit of Fig. 4. It consists of a p-channel transistor as a signal driver, and a n-channel current mirror as a load. The resistor in combination with the power supply V_{DD} provide the external bias current. For the circuit simulations taking into account the radiation effects on the threshold voltage and mobility we have incorporated those effects in the MOS transistors models as discussed in section II.2. Furthermore, we have implemented the adaptive

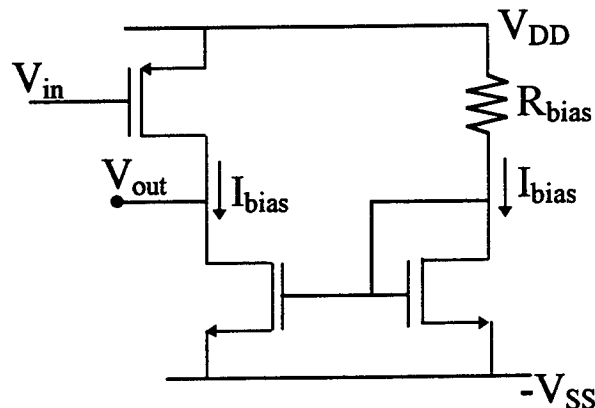


Fig. 7 PMOS Driver Transconductor

scheme to compensate the deviations due to the radiation effects. The implementation of the differential equation of previous section is shown in Fig. 8. There a large number of

potential implementations of the differential equation, the implementation used here involves both operational amplifiers as well as a transconductance amplifier (in these amplifiers the

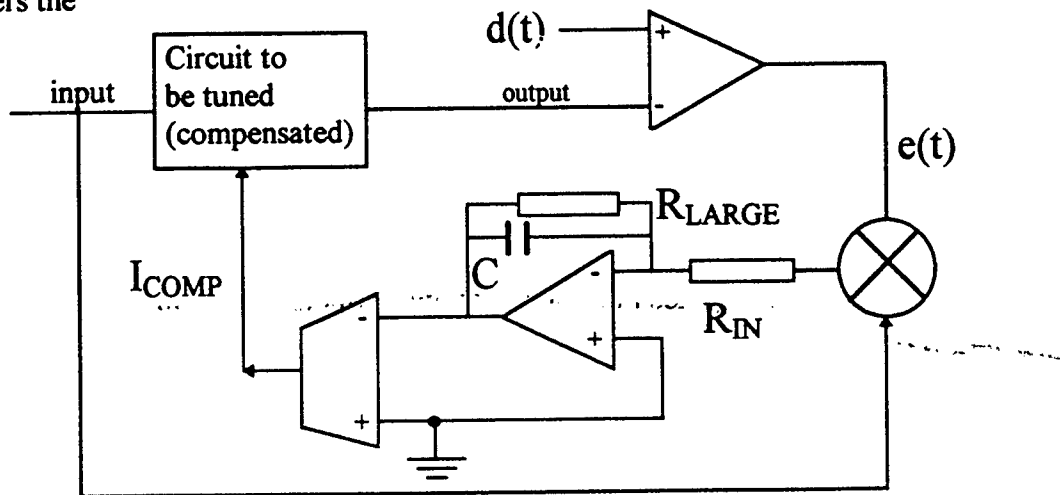


Fig. 8 Macromodel Implementation of (1) for PSPICE

input is voltage and the output is a current). The results for the amplifier for $RAD = 0$ and $10E5$ are shown in Fig. 9(a). It can be observed that a clean sinusoidal drain current (at the top) is obtained under no radiation, also for $RAD = 10E5$ the distorted sinusoidal signal is shown at the bottom. Fig. 9(b) shows the result of using the correction (adaptive) scheme. Both drain

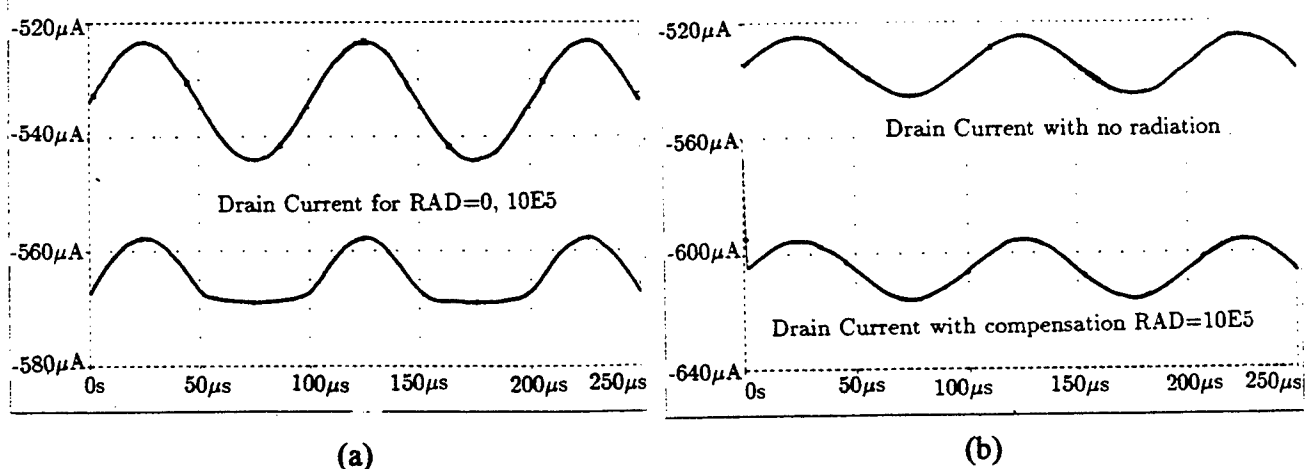


Fig. 9 Drain Current with and without Radiation (a) Regular, (b) Compensated

current curves are illustrated with no-radiation (top trace) and with radiation and compensated. After 2 microseconds the distorted drain current is compensated and a clean sinusoidal current is obtained. Fine details are shown for the output voltage in Fig. 10, in the same figure the ideal output voltage (which is equal to $-30.52 \cdot V_{in}$) is plotted for comparison purposes. It can be seen that after 2 microseconds the adaptive scheme makes the correction, yielding an ideal output signal.

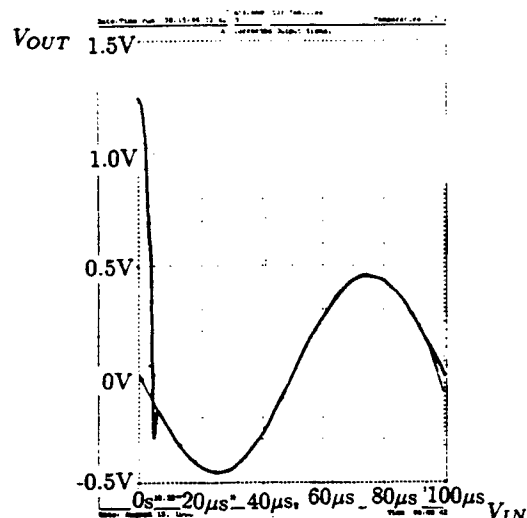


Fig. 10 Output Voltage Compensated and Ideal Output Voltage

VI. CONCLUSIONS AND FURTHER WORK.

The modeling of radiation effects on diodes and MOSFETs using conventional circuit simulators such as PSPICE has been described. Emphasis has been focused on the modeling approach rather than in the actual parameters values, which are process dependent. Relevant issues about low voltage analog circuit design have been discussed. The so called current-mode circuits offer potentially better design, as far as large variations of threshold voltages and mobility, than the conventional voltage-mode circuits. Preliminary results of an adaptive scheme capable to compensate the threshold voltages and mobility variations in a simple transconductor amplifier have been presented.

There are a number of areas that require further research to yield robust analog system tolerant to high radiation effects. Some of these areas are next listed:

(i) Low voltage/ Low power analog circuits [14] tolerant to large parameter process variations. In particular the threshold voltages, mobility and offset (device) currents, and noise [15]. (ii) On-chip testing and fault diagnosis for analog circuits [13]. (iii) Practical adaptive scheme circuit implementations to correct radiation (process variations) effects using conventional process technology such as the ones available from MOSIS. These implementations could be either analog or digital, or a combination depending on the particular application.

References.

- [1] D. R. Alexander, D. G. Mavis, C. P. Brothers, and J. R. Chavez, "Design Issues for Radiation Tolerant Microcircuits in Space," 1996 IEEE Nuclear and Space Radiation Effects Conference , Short Course , pp. V-1 to V-54. NSERC'96. Indian Wells, CA. 1996
- [2] G. C. Messenger, and M. S. Ash, "*The Effects of Radiation on Electronic Systems*," Van Nostrand Reinhold Company, New York, 1986.
- [3] P. Pavan, R. H. Tu, E.R. Minami, G. Lum, P. K. Ko, and C. Hu, " A Complete Radiation Reliability Software Simulator," *IEEE Trans. Nuclear Science*, vol. 41, pp. 2619-2630, December 1994.
- [4] C-C. Chen, S-C. Liu, C-C Hsiao, and J-G Hwu, " A Circuit Design for the Improvement of Radiation Hardness in CMOS Digital Circuits", *IEEE Trans. Nuclear Science*, vol. 39, pp. 272-277, April 1992.
- [5] J.L. Wirth, S.C. Rogers, "The Transient Response of Transistors and Diodes to Ionizing Radiation," *IEEE Trans. Nuclear Science*, NS-11, vol. 5, pp. 24-38, Nov. 1964.
- [6] *MicroSim Application Notes* , pp.207-215, version 6.3, April, 1996.
- [7] *Handbook of Modeling for Circuit Analysis Including Radiation Effects*, Air Force Weapons Laboratory, AFWL-TR-79-86, Albuquerque, New Mexico, May 1979.
- [8] E. Sherra Kerns and B. D. Shafer (editors), The design of radiation-hardened Ics for space: a compendium of approaches, "*Proc. of the IEEE*," vol. 76, No. 11, November, 1988.

- [9] T. L. Turflinger, M. V. Davey, B. M. Mappes, "Single Event Effects in Analog-to-Digital Converters: Device Performance and System Impact," *IEEE Trans. on Nuclear Science*, vol. 41, pp. 2187-2194, December 1994.
- [10] L. Blanquart et al, "Study of Proton Radiation Effects on Analog IC Designed for High Energy Physics in a BiCMOS-JFET RADHARD SOI Technology," *IEEE Trans. on Nuclear Science*, vol. 41, pp. 2525-2529, December 1994.
- [11] R. Koga, et al. "Observation of Single Event Upsets in Analog Microcircuits," *IEEE Trans. on Nuclear Science*, vol. 40, pp. 1838-1844, December 1993.
- [12] H. A. Castro, and M. R. Sweet, "Radiation Exposure Effects on the Performance of an Electrically Trainable Artificial Neural Network (ETANN)," *IEEE Trans. on Nuclear Science*, vol. 40, pp. 1575-1583, December 1993.
- [13] M. Nicolaidis, F. Vargas, and B. Courtois, "Design of Built-In Current Sensors for Concurrent Checking in Radiation Environments," *IEEE Trans. on Nuclear Science*, vol. 40, pp. 1584-1590, December 1993.
- [14] A. Rodríguez-Vázquez, and E. Sánchez-Sinencio, Guest Editors, *IEEE Trans. Circuits and Systems I*, Special Issue on Low-Voltage and Low-Power Analog and Mixed-Signal Circuits and Systems, vol. 43, No. 11, November 1995.
- [15] H. W. Ott, "*Noise Reduction Techniques in Electronic Systems*," John Wiley & sons, New York, 1976.
- [16] S. L. Smith, and E. Sánchez-Sinencio, "Low Voltage Integrators for High-Frequency CMOS Filters Using Current-Mode Techniques," *IEEE Trans. Circuits and Systems II*, vol. 43, pp. 39-48, January 1996.

Appendix 1 : LMS Algorithm Derivation

The mean square error (MSE) is defined as

$$E(t) = 0.5[e(t)]^2 = 0.5[d(t) - y(t)]^2$$

where $d(t)$ is the desired output signal, and $y(t)$ is the actual output signal. The steepest descent algorithm is defined as:

$$\begin{aligned}\frac{dW}{dt} &= -\mu \frac{\partial E}{\partial W} \\ \frac{dW}{dt} &= -\mu \frac{\partial E}{\partial y} \frac{\partial y}{\partial W} \\ \frac{dW}{dt} &= -\mu \frac{\partial [0.5\{d(t) - y(t)\}^2]}{\partial y} \frac{\partial y}{\partial W} \\ \frac{dW}{dt} &= \mu [d(t) - y(t)] \frac{\partial y(t)}{\partial W} \\ \dot{W} &= \mu [d(t) - y(t)] G(t) = \mu e(t) G(t)\end{aligned}$$

Linear System case.

$$y(t) = \sum_{i=0}^n W_i X_i,$$

where:

X_i is the input signal.

Therefore:

$$\frac{dW}{dt} = \mu [d(t) - y(t)] \frac{\partial \sum_{i=0}^n W_i X_i}{\partial W} = \mu [d(t) - y(t)] X_i$$

$$\dot{W} = \mu e(t) X_i$$

Observe that for this case

$$G(t) = X_i(t).$$

Smart Structure/Actuator Modeling and Design for the Integrated Ground
Demonstration Lab

Joseph C. Slater
Assistant Professor
Department of Mechanical and Materials Engineering

Wright State University
3640 Colonel Glenn Highway
Department of Mechanical and Materials Engineering
Dayton, OH 45435
jslater@valhalla.cs.wright.edu

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

September 1996

Smart Structure/Actuator Modeling and Design for the Integrated Ground Demonstration Lab

Joseph C. Slater

1.0 Abstract

The modeling of Ultralight sparse array systems using IMOS (Integrated Modeling of Optical Systems) was studied. Techniques for maintaining model accuracy while decreasing model size (reducing model degrees of freedom) are analyzed. Code has been generated to supplement the capabilities of IMOS. Techniques for the coupling of state space structural systems (systems with common states) in MATLAB were developed. Code for obtaining the observability and controllability matrices of a very large state space system quickly (approximately 350 times faster than the standard MATLAB code) were developed.

Smart Structure/Actuator Modeling and Design for the Integrated Ground Demonstration Lab

Joseph C. Slater

2.0 Introduction

Integrated design of systems is becoming an ever more complex task as technical complexity results in greater division of design tasks. Each revision of a subsystem results in a variation in plant behavior, requiring the redesign of other subcomponents. As a result, combining the results of each design task in a seamless fashion for structural evaluation and control design is becoming a vital tool in efficient system design. The capabilities of IMOS are evaluated and expanded where necessary. The resulting structural models are much smaller, and therefore easier to design control for.

3.0 Guyan Reduction

Consider the second order linear matrix model of a dynamic system

$$\begin{aligned} M\ddot{\mathbf{x}} + D\dot{\mathbf{x}} + K\mathbf{x} &= F\mathbf{u} \\ \mathbf{y} &= C_1\mathbf{x} + C_2\dot{\mathbf{x}} \end{aligned} \quad (1)$$

Next, partition the undamped equation of motion by “master” coordinates, \mathbf{x}_1 , and “slave” coordinates, \mathbf{x}_2 .

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_1 \\ \ddot{\mathbf{x}}_2 \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \mathbf{u} \quad (2)$$

The natural frequencies of the system are given by the Rayleigh quotient to be

$$\omega_i^2 = \frac{\Phi_i^T K \Phi_i}{\Phi_i^T M \Phi_i} = \frac{\Phi_{i1}^T M_{11} \Phi_{i1} + \Phi_{i1}^T M_{12} \Phi_{i2} + \Phi_{i2}^T M_{21} \Phi_{i1} + \Phi_{i2}^T M_{22} \Phi_{i2}}{\Phi_{i1}^T K_{11} \Phi_{i1} + \Phi_{i1}^T K_{12} \Phi_{i2} + \Phi_{i2}^T K_{21} \Phi_{i1} + \Phi_{i2}^T K_{22} \Phi_{i2}} \quad (3)$$

where

$$\Phi_i = \begin{bmatrix} \Phi_{i1} & \Phi_{i2} \end{bmatrix}^T, i = 1, \dots, N \quad (4)$$

are the mode shapes. First, note that variation of the resulting natural frequencies is a second order effect with respect to variation of (or errors in) the mode shape approximation. (You may recall

that this is the practical aspect of the Rayleigh-Ritz method that renders it useful). Second, let's assume that the contribution of the last three terms is small relative to the contribution of the first term. In any finite element model, the second and third terms are generally small, regardless of the choice of coordinates. This assumption is valid if the diagonal terms of the matrix M_{22} are small relative to the diagonal terms of M_{11} , usually by an order of magnitude. After neglecting the slave terms of the mass matrix, the resulting undamped equations of motion are,

$$\begin{bmatrix} M_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_1 \\ \ddot{\mathbf{x}}_2 \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \mathbf{u} \quad (5)$$

The slave displacements are the found from the lower partition of the equations of motion to be

$$\mathbf{x}_2 = K_{22}^{-1}(F_2 - K_{21}\mathbf{x}_1) \quad (6)$$

If the matrix F_2 is restricted to have only zero elements through prudent selection of the slave coordinates (or if F_2 is small relative to $K_{21}\mathbf{x}_1$), then a coordinate transformation matrix P can be defined such that

$$\mathbf{x} = P\mathbf{x}_1 = \begin{bmatrix} I \\ -K_{22}^{-1}K_{21} \end{bmatrix} \mathbf{x}_1 \quad (7)$$

Substituting equation (7) into the full equations of motion yields

$$M_r \ddot{\mathbf{x}}_1 + D_r \dot{\mathbf{x}}_1 + K_r \mathbf{x}_1 = F_r \mathbf{u} \quad (8)$$

with the output equation

$$\mathbf{y} = C_1 P \mathbf{x}_1 + C_2 P \dot{\mathbf{x}}_1 \quad (9)$$

where

$$M_r = P^T M P, D_r = P^T D P, K_r = P^T K P, F_r = P^T F \quad (10)$$

Note that the lower partition (or slave partition) is only limited to being small, not zero. Thus forces can be applied to the structure at points removed from the structural model. Also note that the matrices M_{21} and M_{22} are not neglected. Although the coordinate transformation is derived based on the assumption of a small impact by these matrices, the assumption is small relative to

K_{21} and K_{22} , not absolute magnitude of the matrices. Thus it is common for the ratios m_{ii}/k_{ii} to be compared for all i and ranked in descending order. Here i is the coordinate number and m_{ii} and k_{ii} are elements of the mass and stiffness matrices. Those coordinates for which the ratio is smallest are then defined to be the slave degrees of freedom.

The transformation of equation (7) is a constraint on the system (something like a numerical “boundary condition”). Thus the natural frequencies of the resulting reduced model will be greater than or equal to the natural frequencies of the full model. The natural frequencies will intermesh as described by Meirovitch² via the min-max rule. That is, properly implemented, Guyan reduction will result in slightly increased resulting natural frequencies, with the higher frequencies exhibiting more error.

4.0 Eigenvalue Solution of Second Order Linear Matrix Equations

Consider a system of undamped, unforced, second order linear equations of motion of the form

$$M\ddot{\mathbf{x}} + K\mathbf{x} = \mathbf{0} \quad (11)$$

The first method generally presented for solving these equations is to first assume a solution of the form

$$\mathbf{x}(t) = \Phi \sin(\omega t) \quad (12)$$

Substituting into the equation of motion and factoring out $\sin(\omega t)$ yields

$$(K - M\omega^2)\Phi = \mathbf{0} \quad (13)$$

Premultiplying by the inverse of M results in the traditional eigenvalue problem

$$(M^{-1}K - \omega^2 I)\Phi = \mathbf{0} \quad (14)$$

Although this method works well for small systems, many difficulties arise when working with larger systems. First, the inverse of a large nondiagonal matrix is computationally slow and prone to errors. In addition, the matrix $M^{-1}K$ is not symmetric (the product of two symmetric matrices is not symmetric in general). Thus tools for solving symmetric eigenvalue problems are inapplicable.

A better way of determining the natural frequencies and mode shapes is to start by decomposing the mass matrix using Cholesky decomposition such that

$$R^T R = M \quad (15)$$

where R is an upper triangular matrix. Note that the Cholesky decomposition of a real symmetric matrix is algebraically simple, thus it is not prone to significant numerical errors. Making the substitution

$$\mathbf{x}(t) = R^{-1} \mathbf{v}(t) \quad (16)$$

and premultiplying the equation of motion by $(R')^{-1}$ yields

$$I\ddot{\mathbf{v}} + \tilde{K}\mathbf{v} = \mathbf{0} \quad (17)$$

where $\tilde{K} = (R^T)^{-1} K R^{-1}$. Here we can take advantage of the fact that inverse of R' is the transpose of the inverse of R . It should also be noted that the inverse of an upper triangular matrix is algebraically quite simple. The eigenvector problem is now reduced to

$$(\tilde{K} - \omega^2 I)\mathbf{v} = \mathbf{0} \quad (18)$$

which can easily be solved for significantly large matrices using standard eigensolvers. The mode shapes are then obtained using the coordinate transformation of equation (16).

5.0 Coupling of State Space Systems with Overlapping States

Consider two mechanical systems connected at some common boundary. The state space equations for each system are given by

$$\begin{bmatrix} \dot{\mathbf{z}}_a \\ \dot{\mathbf{z}}_c \end{bmatrix} = \underbrace{\begin{bmatrix} A_{aa} & A_{ac} \\ A_{ca} & A_{cc} \end{bmatrix}}_{A_1} \begin{bmatrix} \mathbf{z}_a \\ \mathbf{z}_c \end{bmatrix} + \underbrace{\begin{bmatrix} B_a \\ B_c \end{bmatrix}}_{B_1} \mathbf{u}_1 + \begin{bmatrix} 0 \\ \mathbf{f}_c \end{bmatrix} \quad (19)$$

$$\mathbf{y}_1 = \underbrace{\begin{bmatrix} C_a & C_c \end{bmatrix}}_{C_1} \begin{bmatrix} \mathbf{z}_a \\ \mathbf{z}_c \end{bmatrix} + D_1 \mathbf{u}_1 \quad (20)$$

and

$$\begin{bmatrix} \dot{\mathbf{z}}_b \\ \dot{\mathbf{z}}_{c'} \end{bmatrix} = \underbrace{\begin{bmatrix} A_{bb} & A_{bc'} \\ A_{c'b} & A_{c'c'} \end{bmatrix}}_{A_2} \begin{bmatrix} \mathbf{z}_b \\ \mathbf{z}_{c'} \end{bmatrix} + \underbrace{\begin{bmatrix} B_b \\ B_{c'} \end{bmatrix}}_{B_2} \mathbf{u}_2 + \begin{bmatrix} 0 \\ \mathbf{f}_{c'} \end{bmatrix} \quad (21)$$

$$\mathbf{y}_2 = \underbrace{\begin{bmatrix} C_b & C_{c'} \end{bmatrix}}_{C_2} \begin{bmatrix} \mathbf{z}_b \\ \mathbf{z}_{c'} \end{bmatrix} + D_2 \mathbf{u}_2 \quad (22)$$

Here the indices c and c' represent the states common to both systems. The vectors \mathbf{f}_c and $\mathbf{f}_{c'}$ represent the equal and opposite forces between the two systems ($\mathbf{f}_c = -\mathbf{f}_{c'}$). Let's assume that the coordinates $\mathbf{z}_{c'}$ are slave coordinates to \mathbf{z}_c . That is, the states $\mathbf{z}_{c'}$ are obtained by setting them equal to the states \mathbf{z}_c as determined from the equations of motion of the first system. Thus the equations for the second system will be used to obtain \mathbf{z}_b and $\mathbf{f}_{c'}$. Partitioning and rewriting the state equations for system two yields

$$\dot{\mathbf{z}}_b = A_{bb}\mathbf{z}_b + \underbrace{\begin{bmatrix} A_{bc'} & B_b \end{bmatrix}}_{B_2'} \begin{bmatrix} \mathbf{z}_{c'} \\ \mathbf{u}_2 \end{bmatrix} \quad (23)$$

and

$$\mathbf{y}_2 = C_b\mathbf{z}_b + \underbrace{\begin{bmatrix} C_{c'} & D_2 \end{bmatrix}}_{D_2'} \begin{bmatrix} \mathbf{z}_{c'} \\ \mathbf{u}_2 \end{bmatrix} \quad (24)$$

$$\mathbf{f}_{c'} = \dot{\mathbf{z}}_{c'} - A_{c'b}\mathbf{z}_b - A_{c'c'}\mathbf{z}_{c'} - B_{c'}\mathbf{u}_2 \quad (25)$$

To be in standard state space form, the equations of motion for the first system must be modified to allow for the additional pseudo-input. The first system is then modified so that the matrix B and vector \mathbf{u}_1 become

$$B_1' = \begin{bmatrix} B_a & 0 \\ B_c & I \end{bmatrix}, \quad \mathbf{u}_1' = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{f}_c \end{bmatrix} \quad (26)$$

The additional outputs to the second system are provided for by modification of the output equation as follows

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{z}_c \\ \dot{\mathbf{z}}_c \end{bmatrix} = \underbrace{\begin{bmatrix} C_a & C_c \\ 0 & I \\ A_{ca} & A_{cc} \end{bmatrix}}_{C_1'} \begin{bmatrix} \mathbf{z}_a \\ \mathbf{z}_c \end{bmatrix} + \underbrace{\begin{bmatrix} D_1 & 0 \\ 0 & 0 \\ B_c & I \end{bmatrix}}_{D_1'} \mathbf{u}_1' \quad (27)$$

Figures 1, 2, and 3 illustrate how the two subsystems are linked using a block diagram.

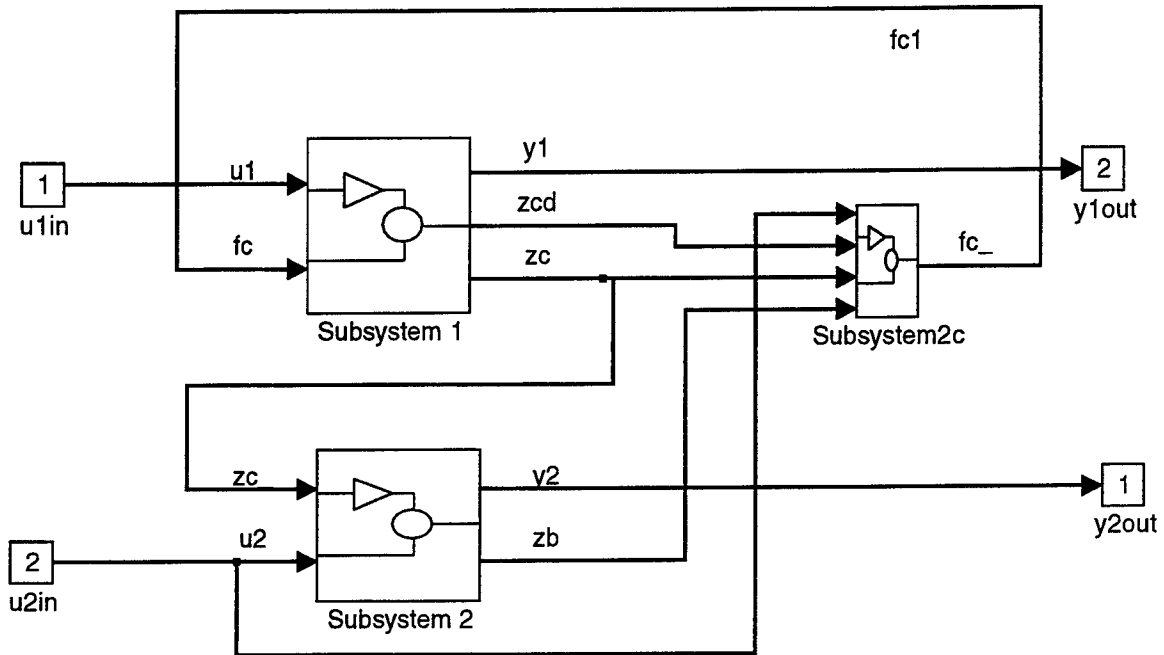


Figure 1: Simulink block diagram linking two modified state space systems.

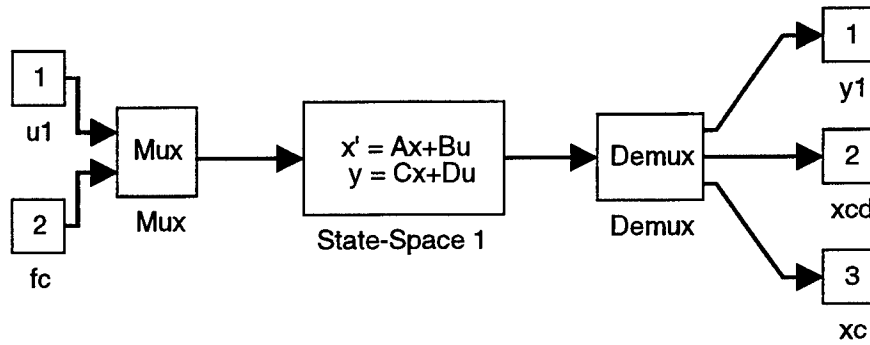


Figure 2: Subsystem 1

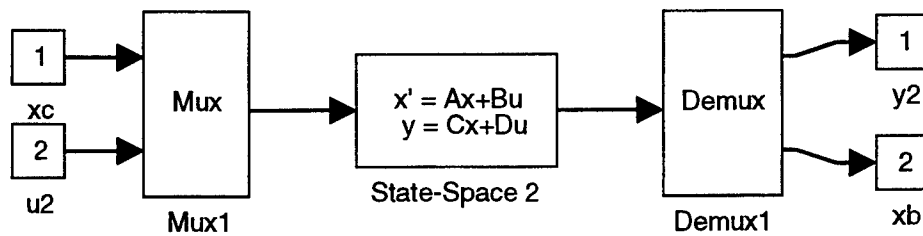


Figure 3: Subsystem 2

6.0 Coupling of Modal Form State Space Systems With Common States

Consider the case where one or both of the models are obtained in modal form such that the state space equations are given by

$$\dot{\mathbf{z}}_m = \mathbf{A}_m \mathbf{z}_m \quad (28)$$

where

$$A_m = \begin{bmatrix} 0 & I \\ -\Lambda_k & -\Lambda_c \end{bmatrix}, \quad (29)$$

$\Lambda_k = \text{diag}(\omega_i^2)$ is a diagonal matrix of the natural frequencies squared, and $\Lambda_c = \text{diag}(2\zeta_i\omega_i)$ is a diagonal matrix of the natural frequencies times the corresponding modal damping ratios times 2. Slater⁴ has shown that the assumption of modal damping is exact when a structure is made of a single material, or when all structural deformation occurs in materials with identical loss factors. Consider the case where the given mode shape matrix, Φ , satisfies the conditions

$$\Phi^T M \Phi = I \quad (30)$$

and

$$\Phi^T K \Phi = \Lambda_k \quad (31)$$

Recall now that the matrix Φ is used in coordinate system transformation between modal and physical coordinates such that

$$\mathbf{x} = \Phi \mathbf{r}, \quad \dot{\mathbf{x}} = \Phi \dot{\mathbf{r}} \quad (32)$$

where \mathbf{x} is the physical coordinate system displacement vector, \mathbf{r} is the modal coordinate displacement vector, and $(\dot{})$ represents a derivative with respect to time. The linear coordinate transformation from the modal state space coordinate system, \mathbf{z}_m to the sorted (common and non-common coordinates) physical state space system \mathbf{z} is given by

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_a \\ \mathbf{z}_c \end{bmatrix} = \begin{bmatrix} \mathbf{x}_a \\ \dot{\mathbf{x}}_a \\ \mathbf{x}_c \\ \dot{\mathbf{x}}_c \end{bmatrix} = \begin{bmatrix} \Phi_a & 0 \\ 0 & \Phi_a \\ \Phi_c & 0 \\ 0 & \Phi_c \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \dot{\mathbf{r}} \end{bmatrix} = \Phi_{ss} \mathbf{z}_m \quad (33)$$

Note that if the number of states kept in the physical state space system is equal to the number of modes, then the matrix Φ_{ss} is square and directly invertible. If the number of physical degrees of freedom kept is less than the number of modes retained in the model, additional modal truncation must be performed. In the case where more coordinates are retained in the physical coordinate system than there are modes in the modal model, only a subset of the retained degrees of

freedom can be used in the physical space equations of motion. The remaining degrees of freedom can then be written as a linear combination of the coordinates in the state space equation. The issue of a non-square modal matrix is not addressed here. However, the derivations are relatively straight forward if needed.

Consider the inverse coordinate transformation

$$\mathbf{z}_m = \Phi_{ss}^{-1} \mathbf{z} \quad (34)$$

Substituting this coordinate transformation into the state space modal equation of motion (28) yields

$$\Phi_{ss}^{-1} \dot{\mathbf{z}} = \begin{bmatrix} 0 & I \\ -\Lambda_k & -\Lambda_C \end{bmatrix} \Phi_{ss}^{-1} \mathbf{z} \quad (35)$$

Premultiplying by Φ_{ss} yields the state space equations of motion in physical coordinates (19)

$$\dot{\mathbf{z}} = \Phi_{ss} \begin{bmatrix} 0 & I \\ -\Lambda_k & -\Lambda_C \end{bmatrix} \Phi_{ss}^{-1} \mathbf{z} = A_1 \mathbf{z} \quad (36)$$

It is imperative that when performing the computation to obtain A_1 in MATLAB that the forward slash operator, /, be used instead of multiplication by an inverse. Using the forward slash operator forces MATLAB to perform LU decomposition on the matrix Φ_{ss} and subsequently perform Gaussian elimination to obtain the result of $A_m \Phi_{ss}^{-1}$. Bypassing the explicit inverse of Φ_{ss} results in faster and more accurate computation of the state matrix.

7.0 Rigid Beam Elements

When a beam in a structure is sufficiently stiff such that very little deformation occurs relative to other members of the structure, it is often possible to simply assume that the member is rigid. When a beam is rigid, the motion of the degrees of freedom are linked in such a fashion that the motion of one node is directly dependent on the motion of the other. In similar fashion to Guyan reduction we define a master and a slave coordinate. In the following, node 1 is defined at the master element and node 2 is defined as the slave element. Consider the following rigid beam element:

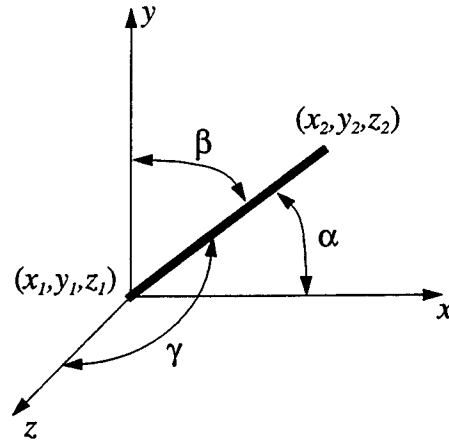


Figure 4: Beam in three dimensional Cartesian coordinates.

Here node one is placed at the local coordinate system origin and the angles α , β , and γ denote the angles between the vector $(x_2 - x_1)\hat{i} + (y_2 - y_1)\hat{j} + (z_2 - z_1)\hat{k}$ and the coordinate axes. The equilibrium positions of nodes one and two are defined by (x_1, y_1, z_1) and (x_2, y_2, z_2) . Degrees of freedom of the nodes are shown in Figure 5.

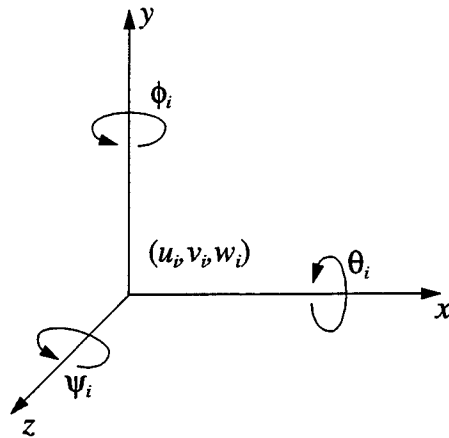


Figure 5: Degrees of freedom of node i .

Consider for the moment that all rotations are zero. Then the displacements (u_2, v_2, w_2) are equal to the displacements (u_1, v_1, w_1) . In addition, the translation of node 1 without rotation does not cause a rotation of node 2. Next consider the case that all translations are zero. Then the rotations of node 2 are equal to the rotations of node 1. However, the translations of node 2 depend on the orientation and geometry of the beam (or the relative positions of the two nodes).

To derive these displacements, let's first pose the problem as one of obtaining the elements of a matrix A such that

$$\begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = A \begin{bmatrix} \theta_1 \\ \phi_1 \\ \psi_1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} \theta_1 \\ \phi_1 \\ \psi_1 \end{bmatrix} \quad (37)$$

First consider rotation only about the x axis (θ_1). The first step is to project the vector onto the yz plane, since no displacement will take place in the x direction ($a = 0$).

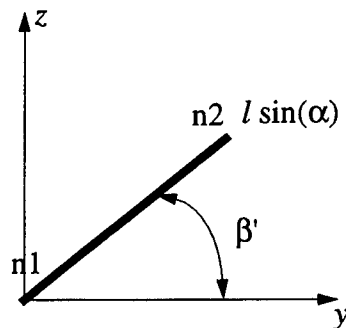


Figure 6: Rigid beam in yz plane.

From basic trigonometry it can now be seen that a small rotation θ_1 about the x axis will result in a translation of $-\theta_1 l \sin(\alpha) \sin(\beta')$ in the y direction and a translation of $\theta_1 l \sin(\alpha) \cos(\beta')$ in the z direction.

Next consider rotation only about the y axis (ϕ_1). The first step is to project the vector onto the xz plane, since no displacement will take place in the x direction ($e = 0$).

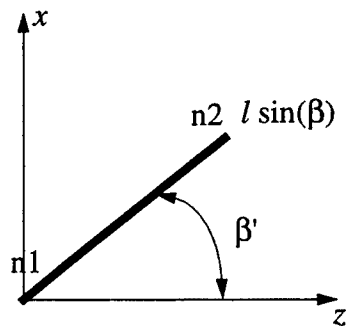


Figure 7: Rigid beam in xz plane.

A small rotation ϕ_i about the x axis will result in a translation of $-\phi_i l \sin(\beta) \sin(\gamma')$ in the z direction and a translation of $\phi_i l \sin(\beta) \cos(\gamma')$ in the x direction.

Finally, consider rotation only about the z axis (ψ_i). The first step is to project the vector onto the xy plane, since no displacement will take place in the z direction ($i = 0$).

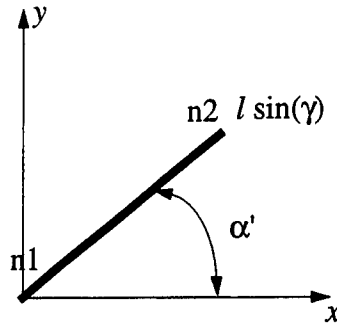


Figure 8: Rigid beam in xy plane.

Likewise, a small rotation ψ_i about the z axis will result in a translation of $-\psi_i l \sin(\gamma) \sin(\alpha')$ in the x direction and a translation of $\psi_i l \sin(\gamma) \cos(\alpha')$ in the y direction.

Combining the results of the preceding analyses results in the coordinate transformation

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \\ \theta_1 \\ \phi_1 \\ \psi_1 \\ u_2 \\ v_2 \\ w_2 \\ \phi_2 \\ \phi_2 \\ \psi_2 \end{bmatrix} = T \begin{bmatrix} u_1 \\ v_1 \\ w_1 \\ \theta_1 \\ \phi_1 \\ \psi_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \\ I & A \\ 0 & I \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ w_1 \\ \theta_1 \\ \phi_1 \\ \psi_1 \end{bmatrix} \quad (38)$$

where

$$A = \begin{bmatrix} 0 & l \sin \beta \cos \gamma' & -l \sin \gamma \sin \alpha' \\ -l \sin \alpha \sin \beta' & 0 & l \sin \gamma \cos \alpha' \\ l \sin \alpha \cos \beta' & -l \sin \beta \sin \gamma' & 0 \end{bmatrix} \quad (39)$$

The resulting matrix T can then be used as a coordinate transformation matrix so that the resulting second order system matrices become

$$M_r = T^T M T, D_r = T^T D T, K_r = T^T K T, F_r = T^T F \quad (40)$$

and the resulting output equation is then

$$y = C_1 T x_1 + C_2 T x_1 \quad (41)$$

Codes for obtaining the matrices M_r , K_r , and T can be found in Appendices 10.1 and 10.2.

8.0 Testing of Piezo-Actuators at Very Low Amplitudes

In application, piezo stack actuators are used in one of four configurations, open loop or closed loop, and force control or position control. In open loop control, some assumption must be made about the boundary conditions of the actuator so that the proper control voltage may be determined, from the constitutive laws, and applied. Consider the spring equation (derived from the PZT constitutive laws) for a PZT cylinder in extension:

$$\Delta L = d_{33} L_0 E + \frac{1}{K} F \quad (42)$$

where ΔL (m) is the change in length, d_{33} (m/V) is the piezoelectric constant in the 3-3 direction (the 3 direction is the direction of polarization, assumed to be along the cylinder), L_0 (m) is the length of the cylinder, E (V/m) is the total electric field in the material, K (N/m) is the stiffness of the cylinder, and F (N) is the externally applied load (equal and opposite to the load applied by the PZT to the mounts). That is, depending on the reaction force, F , and the electric field, F , in the material, a displacement ΔL will occur.

This sounds pretty simple... too simple. Unfortunately it is when attempting high resolution control of a PZT actuator. Any real material inevitably has some damping resulting in hysteresis. This is usually modeled as a complex stiffness, $K^* = K + K' i$ where $i = \sqrt{-1}$ and K' is referred to as the loss modulus. The complex stiffness may also be written as $K^* = K(1 + \eta i)$ where $\eta = K'/K$ is called the material loss factor. Certain restrictions apply to these constants for com-

patibility sake.³ These values are often frequency and material dependent, and therefore can be difficult to model with a great degree of accuracy.⁴

In addition, the total electric field is not the same as the applied electric field. The total electric field in a dielectric is a function of both the applied electric field and the electric field resulting from material polarization. However, the material polarization is a function of the total electric field, this interdependence results in some creep in the response of the PZT. Main, Garcia, and Newton¹ have shown that the first order effects of this phenomenon can be compensated for, resulting in an greatly improved response (not quantified by the authors). I am not aware of any model of the delay that causes the hysteresis. It is possible that electrical models similar to the mechanical models used by Slater et. al.⁴ may be a viable.

9.0 References

¹Main, J. A., Garcia, E., and Newton, D. V., "Precision position control of piezoelectric actuators using charge feedback," *Journal of Guidance, Control, and Dynamics*, v. 18, Sept./Oct '95, pp. 1068-73.

²Meirovitch, L., *Computational Methods in Structural Dynamics*, Sijthoff & Noordhoff, Rockville, Md., U.S.A., 1980

³Nashif, A. D., Jones, D. I. G., and Henderson, J. P., *Vibration Damping*, Wiley, New York, 1985.

⁴Slater, J.C., Belvin, W.K., and Inman, D.J., "A Survey of Modern Methods for Modeling Frequency Dependent Damping in Finite Element Models," *Proceedings of the 11th International Modal Analysis Conference*, 1993.

10.0 Appendix

10.1 The following is a MATLAB code for obtaining the local rigid coordinate transformation matrix to rigidly connect one node to another in a frame finite element model.

```
function [T]=rbeam(n1,n2)

% RBEAM Rigid beam element.
% [T]=RBEAM(n1,n2) generates coordinate transformation matrix for
% a rigid beam element such that the reduced mass matrix is
% MR=T'*M*T and the reduced stiffness matrix is KR=T'*K*T
% where M and K are the submatrices of the global matrices
% corresponding to the nodes defined by n1 and n2.
% n2 is fixed to n1. The row vectors n1 and n2 are
% n1=[node1#, node1x, node1y, node1z;...]
% n2=[node2#, node2x, node2y, node2z;...]
% where the second through fourth columns are node locations.
```

```

% Version 1.0
% Copyright Joseph C.Slater 8/28/96

x1=n1(2);
y1=n1(3);
z1=n1(4);
x2=n2(2);
y2=n2(3);
z2=n2(4);

l=sqrt((x2-x1)^2+(y2-y1)^2+(z2-z1)^2);
alphap=atan2(y2-y1,x2-x1);
betap=atan2(z2-z1,y2-y1);
gammap=atan2(x2-x1,z2-z1);
pv=[x2-x1 y2-y1 z2-z1];
lam=pv/norm(pv);
alpha=acos(lam(1));
beta=acos(lam(2));
gamma=acos(lam(3));

a=0;
b=l*sin(beta)*cos(gammap);
c=-l*sin(gamma)*sin(alphap);
d=-l*sin(alpha)*sin(betap);
e=0;
f=l*sin(gamma)*cos(alphap);
g=l*sin(alpha)*cos(betap);
h=-l*sin(beta)*sin(gammap);
i=0;

T=[eye(3) zeros(3,3);...
   zeros(3,3) eye(3);...
   eye(3) [a b c;d e f;g h i] ;...
   zeros(3,3) eye(3)];

```

10.2 The following is a MATLAB code for obtaining the global rigid coordinate transformation matrix rigidize multiple frame members in a finite element model.

```

function [mr,kr,T,pout]=rigid(m,k,no1,no2,pin)

% RIGID Assemble rigid beam elements into mass and stiffness matrices.
% [Mr,Kr,T,pout]=RIGID(M,K,n1,n2,pin) generates coordinate
% transformation matrix for forming rigid connections between
% nodes in matrices n1 and n2. The matrices are reduced via
% Mr=T'*M*T and Kr=T'*K*T where M and K are the global matrices.
% The matrices n1 and n2 are pairs of row vectors in the form
% n1=[node1#, node1x, node1y, node1z;...]
% n2=[node2#, node2x, node2y, node2z;...]
% n2 represents the DOFs to be removed from the system.

```

```

%      pin is a vector of DOF numbers corresponding to the row/columns
%      of M and K. pout is the vector of DOF numbers after the matrices have
%      been reduced. (pin with the nodes n2 removed)
%
%      Note that in version 1 the same node may not appear in both
%      n1 and n2. This does not result in a loss of generality, but
%      only in slight inconvenience.

% Version 1.1
% Copyright Joseph C. Slater 7/31/96
% Revised 8/28/96, JCS, test for input errors.

[rn1,cn1]=size(no1);
[rn2,cn2]=size(no2);

% Check sizes of inputs
if rn1 ~= rn2
    disp('n1 and n2 must be the same size.')
    break
end

% Check to make sure no nodes numbers are common to n1 and n2
mn=max([no1(:,1); no2(:,1)]);
a=zeros(mn,1);a=sparse(a);b=a;
a(no1(:,1))=ones(rn1,1);
b(no2(:,1))=ones(rn2,1);
c=a+b;
if max(c)>1
    disp(' ')
    disp(setstr(7))
    bn=find(c==2);
    lbn=length(bn);

    if lbn>1
        w1='Nodes ';
        w2=' have ';
        bnsf=[' and ' num2str(bn(lbn))];
    else
        w1='Node ';
        w2=' has ';
        bnsf='';
    end
    bns=num2str(bn(1));
    for i=2:length(bn)-1
        bns=[bns ' ' num2str(bn(i)) ];
    end
    bns=[bns bnsf];

    disp([w1 bns w2 'been found to be common to both n1 and n2.'])
    disp('This is bad, please don''t do it again.')
    disp(' ')

```

```

        break
    end

    sp=length(k);

    if nargin==4
        pin=1:sp;
    end

    pin=pin(:);
    p=pin;
    T=eye(sp);

    [a,i]=sort(-no2(:,1));

    n1=no1(i,:);
    n2=no2(i,:);
    [n,a]=size(no1);

    for i=1:n

        t=rbeam(n1(i,:),n2(i,:));
        [r,c]=size(T);

        l1=find(pin==n1(i,1));
        l2=find(pin==n2(i,1));

        T=[T(:,(1:l1-1)*6) T(:,n1*6+1:c)];
        T((n2-1)*6+1:n2*6,(n1-1)*6+1:n1*6)=t(7:12,:);
        p=[p(1:(n2-1)*6);p(n2*6+1:c)];
    end

    % Remaining degrees of freedom
    pout=p;

    % Reduced matrices
    mr=T'*m*T;
    kr=T'*k*T;

```

10.3 The following is a MATLAB code for obtaining the observability matrix of a state space system. For a 240 state system the code is approximately 350 times faster than the code provided in the MATLAB Control Toolbox (seconds as opposed to hours).

```

function ob = obsv2(a,c)
%OBSV2  Forms observability matrix.
%      OBSV2 uses a much more efficient algorithm
%      than OBSV, reducing the number of matrix
%      multiplications from ((N-1)^2)/2 to N-1.
%      For large systems the speed improvement is multiple
%      orders of magnitude.

```

```

%      Ob = OBSV2(A,C) returns the controllability matrix
%      Co = [C ; CA ; CA^2 ; ... ; CA^(N-1)]
%
%      See also: OBSV, OBSVF.

%      Authors : Joseph C. Slater and Jesse A. Leitner 1996

[p,n]=size(c);
ob = zeros(n*p,n);
nn=1:n;
pp=1:p;

ob(pp,nn)=b;

for i=1:n-1
    ob(p*i+pp,:)=ob(p*(i-1)+mm,:)*a;
end

```

10.4 The following is a MATLAB code for obtaining the controllability matrix of a state space system. For a 240 state system the code is approximately 350 times faster than the code provided in the MATLAB Control Toolbox (seconds as opposed to hours).

```

%CTRB2  Forms controllability matrix.
%      CTRB2 uses a much more efficient algorithm
%      than CTRB, reducing the number of matrix
%      multiplications from ((N-1)^2)/2 to N-1.
%      For large systems the speed improvement is multiple
%      orders of magnitude.
%      Co = CTRB2(A,B) returns the controllability matrix
%      Co = [B AB A^2B ... A^(N-1)B]
%
%      See also: CTRB, CTRBF.

%      Authors : Joseph C. Slater and Jesse A. Leitner 1996
%      Copyright (c) 1996 by Joseph C. Slater and Jesse A. Leitner.

[n,m]=size(b);
co = zeros(n,n*m);
nn=1:n;
mm=1:m;

co(nn,mm)=b;

for i=1:n-1
    co(:,m*i+mm)=a*co(:,m*(i-1)+mm);
end

```

MODELING OF TOTAL DOSE RESPONSE OF SOI N-MOSFETS FOR LOW POWER CMOS CIRCUITS

Ashok Srivastava
Associate Professor

Department of Electrical and Computer Engineering

Louisiana State University
Baton Rouge, LA 70803-5901

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC
and
Phillips Laboratory
August 1996

CMOS MODELING OF TOTAL DOSE RESPONSE OF SOI N-MOSFETS FOR LOW POWER CMOS CIRCUITS

Ashok Srivastava
Associate Professor
Department of Electrical and Computer Engineering
Louisiana State University

Abstract

In the present work, we have developed an integrated and flexible rad-sensitive sub-threshold model for the partially depleted, SOI n-channel MOSFETs based on existing pre-irradiated device models for applications in low power CMOS circuit designs. The radiation-induced oxide trapped charge and interface state charge are integrated in the charge-sheet model of the MOSFET. The model is continuous from subthreshold to strong inversion regions, with floating body effects, and associated parasitics such as BJT included. To demonstrate the applicability of the model, we have predicted subthreshold current-voltage characteristics of a typical SOI n-MOSFET under a total dose condition in linear region of shift in threshold voltage versus dose. The predicted shift of subthreshold characteristics to more negative values with dose with respect to pre-irradiated condition agrees with the experimentally observed behavior reported in literature.

MODELING OF TOTAL DOSE RESPONSE OF SOI N-MOSFETS FOR LOW POWER CMOS CIRCUITS

Ashok Srivastava

Introduction

SOI/CMOS circuits offer a higher level of performance in speed, and power dissipation due to reduced parasitic junction capacitances {1, 2} and radiation hardness due to the small volume of SOI film {3} compared to bulk CMOS circuits. The other advantages of SOI/CMOS are absence of latchup and applications in high-temperature electronics {3}. The major applications where the SOI devices and circuits are used are the aerospace and military fields {4}. These applications have resulted in significant theoretical and experimental research on SOI devices and circuits in simulated radiation environments for space electronics {5-8}. It has been observed that the thicker, partially depleted n-SOI MOSFETs offer superior performances in a harsh environment {3}. These devices are better than or equal to their corresponding bulk devices in total dose hardness and better than bulk devices in SEU hardness.

The basic structure of bulk CMOS and SOI/CMOS devices is same. However, the fundamental difference lies in SOI film thickness, floating body and back-gate {34}. This makes the physics of SOI devices interesting and challenging. Significant research has been done in past on SOI devices modeling by various groups {9-12}. Analytical and CAD models have been developed for SOI/CMOS circuit designs {11-14}. Theories for bulk MOS devices have been extended to SOI devices with floating body effects.

In today's electronics, power dissipation has become as important a consideration as speed and area due to portable and battery operated microelectronic systems in computation and communication areas. In CMOS digital circuits, the dynamic power dissipation is proportional to the square of the power supply voltage. Lowering the supply voltage will require the lowering of the operating speed as high as possible. The logic may switch between subthreshold and near threshold regions {16,17}. This requires proper understanding and modeling of MOSFET's operation in the subthreshold region. In SOI devices, the floating body affects the characteristics including those in the subthreshold region {15}. When SOI devices are operated in space radiation environments, the subthreshold I-V characteristics further degrade as evidenced by increase in subthreshold slope. The degradation increases with the increase in dose levels {6,18,19} and thus affects operation and performance of SOI/CMOS circuits.

At present, the operation of SOI devices in the subthreshold region is the subject of significant experimental research under the total dose environment for the development of better radiation hardened SOI/CMOS circuits. Research papers have been published to model SOI devices using SPICE {13} and numerical techniques {20} under total dose conditions. However, RAD SOI MOSFET models have been focused to bulk-MOS devices; limited only to subthreshold region; confined mainly to non-uniform distribution of interface traps in oxide; neglected relevant charge terms in oxide models; used numerical or commercially available tools such as TMA and SPICE; and neglected parasitics {20,32,33}. In the present work, the investigator has developed an integrated and flexible rad-sensitive model to predict subthreshold behavior of partially depleted, SOI n-channel MOSFETs under low total dose conditions based on existing pre-irradiated device models for low power operation.

Charge Buildup in Oxide and Mathematical Model

Figure 1(a) shows charge buildup and removal processes in a MOS structure. The electric field in the oxide is shown in Fig. 1(b). Electron-hole pairs are generated throughout the oxide by the ionizing radiation. In the presence of electric field, electrons move toward the positive gate and holes toward the substrate. Depending upon the density of traps near the Si-SiO₂ interface and hole capture cross-section of the traps, a fraction of the holes are trapped on trapping centers and the remainder continue to move toward the silicon substrate and are lost due to recombination {21}. Some of the holes captured by hole traps within a tunneling distance of the silicon substrate (~ 2.5-5 nm) are assumed to be rapidly removed by recombination with electron tunneling from the silicon {22}.

It is to be mentioned that the transport, trapping, injection and recombination processes all are functions of the electric field, $E(x)$, across the oxide. The trapped hole distribution modifies the electric field distribution in the oxide as shown in Fig. 1(b). E_0 is the initial value of the electric field. The electric field is enhanced toward the substrate to a value E_2 and reduces toward the gate to a value E_1 with respect to E_0 , due to trapped charge and positive bias.

A simple mathematical model to explain thin sheet of charge produced in a MOS device by hole trapping during irradiation under an electric field is shown in Fig. 2. Q_{ot} is a thin sheet of charge corresponding to hole traps and is located at a distance x_1 and x_2 from the substrate and gate, respectively. Regions 1 and 2 define the charge collection regions. Q_s is the image charge density in the substrate (Si) space charge region. Note that there may be some negative image charge of density, Q_m on the metal gate. For all practical purposes and modeling $Q_m \ll Q_s$. Thus, a

charge of $-Q_s$ must be applied to the gate in order to restore the charge-state of semiconductor corresponding to its pre-irradiation condition. The potential corresponding to this charge is equal to the negative shift in threshold voltage and is given by {23},

$$\Delta V_{th}(Q_{ot}) = -\frac{(qN_{ot})x_2}{\epsilon_o \epsilon_{ox}} \quad (1)$$

where N_{ot} is the number of trapped charge per unit area and q is the electronic charge. ϵ_{ox} and ϵ_o are the dielectric constant of oxide and permittivity of free space, respectively.

The electron-hole recombination process in oxide which follows immediately after the irradiation, reduces the total amount of charges available for transport under the applied electric field. The probability that a hole escapes recombination either reaches the substrate or is trapped in the oxide, can be defined by the probability function $f(E)$. The function $f(E)$ is normally called a 'fractional charge yield'. If the fraction, 'A' is the probability that the hole will be trapped on passing through the trapping sheet, the number of holes trapped per unit area (N_{ot}) is given by {23}

$$N_{ot} = g \cdot x_c \cdot f(E) \cdot A \quad cm^{-2}rad^{-1} \quad (2)$$

The rate of generation of holes in oxide due to incident ionization radiation is defined by 'g' and its value is 7.9×10^{12} holes $cm^{-3} rad^{-1}$ in SiO_2 . The thickness x_c is the charge collection region and is equal to x_2 for positive gate bias. It is to be mentioned that the parameter 'A' in eq. 2 depends upon the area density of traps and hole capture cross-section. The choice of 'A' has a significant influence on N_{ot} . Combining eqs. 1 and 2, we obtain

$$\Delta V_{th} = -\frac{qg}{\epsilon_o \epsilon_{ox}} \cdot x_2 x_c \cdot f(E) \cdot A \cdot D \quad (3)$$

Let

$$R = \frac{qg}{\epsilon_o \epsilon_{ox}} \cdot x_2 x_c \cdot f(E)$$

Equation 3 can be put in the form

$$\Delta V_{th} = RAD$$

D is the dose in rad (SiO₂). The fractional charge yield, $f(E)$ is determined from an auxiliary experiment to measure electron injection and oxide charge generation in MOS devices in a 10 KeV X-ray source using a photocurrent measurement technique. The total photocurrent is a sum of conduction current and injection current. The conduction current includes $f(E)$ and the injection current depends upon the condition near the interface. From a best fit to experimental data{22}, $f(E)$ is of the form

$$f(E) = \left[\frac{1.30}{E + 0.113} + 1 \right]^{-1} \quad (4)$$

By combining eqs. 3 and 4, we obtain

$$\Delta V_{th} = -\frac{qg}{\epsilon_o \epsilon_{ox}} \cdot x_2 x_c \cdot \left[\frac{1.30}{E + 0.113} + 1 \right]^{-1} \cdot A \cdot D \quad (5)$$

We obtain Q_s or Q_{ot} from eq. 5

$$-\frac{Q_s}{q} = \frac{Q_{ot}}{q} = \frac{g}{t_{ox}} \cdot x_2 x_c \cdot \left[\frac{1.30}{E + 0.113} + 1 \right]^{-1} \cdot A \cdot D \quad (6)$$

where t_{ox} is the oxide thickness.

Up to now, we have considered only the effect of positive charge trapping in oxide. However, hole capture in oxide is also accompanied by the creation of chargeable interface states at the Si-SiO₂ interface. Depending on the type of interface states, the interface state charge tends to compensate or enhance the total threshold voltage shift {8}. In current MOS technology, interface states are negatively charged as discussed below.

There are many explanations in literature for generation of interface states due to radiation {21}. It has been shown by Shino et al. {24} that interface trap generation appeared only under a positive oxide field. This indicates that hole transport to the Si-SiO₂ interface is responsible for interface trap generation. The equivalent shift in threshold voltage can be obtained from the equation {25},

$$\Delta V_{th}(Q_{it}) = \frac{q}{\epsilon_o \epsilon_{ox}} \cdot g \cdot f(E) \cdot f_{it} \cdot (t_{ox})^2 \cdot D \quad (7)$$

The corresponding charge, Q_{it} is given by

$$\frac{Q_{it}}{q} = -\frac{g}{t_{ox}} \cdot f(E) \cdot f_{it} \cdot (t_{ox})^2 \cdot D \quad (8)$$

where f_{it} is the interface-trap generation efficiency (i.e., the total number of interface traps created per electron-hole pair).

SOI n-Channel MOSFET Model

Figure 3 shows the cross-section of an n-channel SOI MOSFET. Figure 4 shows the corresponding equivalent circuit which includes the parasitic floating base bipolar transistor and the back-gate MOS transistor. I_D and I_{IC} are the channel and impact ionization currents. I_{BS} and I_{BD} are the diffusion and leakage currents corresponding to body-source and body-drain junction diodes. I_{BB} is the recombination component of the BJT base current, I_B and $I_B \approx I_{BB}$. I_E and I_{CE} are the emitter and collector currents of the BJT. The impact ionization current source due to parasitic MOSFET at its drain is shown by the dotted line. Its effect is not included in the analysis. This is because the normal circuit operation requires the substrate voltage, V_{SUB} is to be kept well below the threshold voltage of the parasitic MOSFET.

Circuit designs whether digital or analog often requires more accurate device models for use in circuit simulation. Circuit performance implemented in SOI/CMOS technology degrades due to floating body effects. Many SOI MOSFET models for use in circuit designs have been reported in literature but are restricted to strong inversion [9]. In this work, we will use the charge-sheet model of the MOSFET of Brews [26] which is continuous from subthreshold (weak inversion) to strong inversion regions. However, we will restrict our analysis to the subthreshold region of operation of the MOSFET. The transistor channel current, I_D applicable in an n-MOSFET with floating body can be described by the equation

$$\begin{aligned} I_D = \mu \cdot \frac{W}{L} \cdot \beta^{-1} \cdot \left\{ C_{ox} (1 + \beta V_{gs} - \beta V_{FB} - \beta V_{bs}) \cdot (\Phi_d - \Phi_s) \right. \\ \left. - \frac{\beta}{2} C_{ox} (\Phi_d^2 - \Phi_s^2) \right. \\ \left. - q N_A L_B \left(\frac{2\sqrt{2}}{3} \right) \cdot [(\beta \Phi_d - 1)^{3/2} - (\beta \Phi_s - 1)^{3/2}] \right. \\ \left. + q N_A L_B \sqrt{2} \cdot [(\beta \Phi_d - 1)^{1/2} - (\beta \Phi_s - 1)^{1/2}] \right\} \end{aligned} \quad (9)$$

where various parameters are defined as follows.

$$L_B = \sqrt{\frac{kT}{q} \frac{\epsilon_s}{qN_A}} - \text{Debye Length}$$

μ - electron mobility in channel

C_{ox} - gate oxide capacitance/area

W - channel width

L - channel length

$V_s = R_s \cdot I_{DS} = 0$ for R_s (source resistance) = 0

$V_d = V_D - R_D I_{DS} = V_D$ for R_D (drain resistance) = 0

$V_{gs} = V_G - V_s = V_G$ for $V_s = 0$

$V_{bs} = V_B - V_s = V_B$ for $V_s = 0$

V_{FB} - flat-band voltage

Q_s - surface potential at the source end

Q_d - surface potential at the drain end

N_A - p-body doping

ϵ_s - permittivity of silicon

N_A in eq. 9 is assumed to be uniformly doped due to relatively thin SOI film.

The surface potentials, ϕ at source and drain ends of the channel region can be calculated from the solution of an one-dimensional Poisson's equation {27} and modified to include effects of the body potential, V_{bs} , and source and drain bias, V_a . The surface potential is given by

$$C_{ox}(V_{gs} - V_{FB} - V_{bs} - \Phi) = \sqrt{\frac{2\epsilon_s q}{\beta}} \cdot \left\{ N_A (\beta\Phi - 1 + e^{-\beta\Phi}) + \frac{n_i^2}{N_A} \cdot [e^{\beta(-V_{bs}-V_a)} \cdot (e^{\beta\Phi} - 1) - \beta\Phi] \right\}^{1/2} \quad (10)$$

where n_i is the intrinsic carrier concentration of Si. $V_a = V_d$ for $\phi = \phi_D$ and $V_a = V_s$ for $\phi = \phi_s$.

The flat-band voltage, V_{FB} in eq. 10 is described by

$$V_{FB} = \Phi_{ms} - \left(\frac{Q_{ss}}{C_{ox}} \right) \quad (11)$$

where ϕ_{ms} is the metal-semiconductor work function difference and Q_{ss} the fixed oxide charge at the Si-SiO₂ interface. In presence of ionizing radiation, the flat-band condition is modified by adding potentials corresponding to trapped positive charge and interface states charge described by eq. 6. The modified V_{FB} is now given by

$$V_{FB} = \Phi_{ms} - \left(\frac{Q_{ss}}{C_{ox}} + \frac{Q_{ot}}{C_{ox}} - \frac{Q_{it}}{C_{ox}} \right) \quad (12)$$

As the channel current (electrons) flows (move) in the high-field region adjacent to the drain, electron-hole pairs are generated by impact ionization {28}. Generated electrons flow with the channel current, to the drain, while the holes move towards the place of lowest potential, the floating body. Because of the floating body, holes are forced to flow in the source junction, there by providing a forward bias, V_{bs} across the junction {3}. The impact ionization current, I_{IC} generated in the high field region due to holes near the drain and represented by a current source in Fig. 4 is given by

$$I_{IC} = (M - 1)I_D + I_{CE} \quad (13)$$

In low-current gain parasitic BJT, I_{CE} approaches to zero. $(M-1)$ is the impact ionization coefficient or multiplication factor. The multiplication factor depends on both the drain and gate voltage {29}

$$(M - 1) = \frac{A_i}{B_i} \cdot [V_{ds} - V_{DSAT}] \cdot \exp \left[-B_i \frac{l_c}{V_{ds} - V_{DSAT}} \right] \quad (14)$$

where l_c is the characteristic length. For thin, partially depleted n-MOSFET, l_c can be expressed as

$$l_c = t_b \cdot \left[\frac{\epsilon_s / t_b}{2C_{ox} \left(1 + \frac{\epsilon_s}{C_{ox} \cdot t_b} \right)} \right] \quad (15)$$

where t_b is the SOI film thickness. A_i and B_i in eq. 15 are impact ionization constants for electrons, assumed to be $1.4 \times 10^6 \text{ cm}^{-1}$ and $2.6 \times 10^6 \text{ V/cm}$, respectively.

V_{DSAT} is given by {30}

$$V_{DSAT} = V_{gs} - V_{FB} + 2\Phi_f + \frac{\epsilon_s q N_A}{C_{ox}^2} \left[1 - \left\{ 1 + \frac{2C_{ox}^2}{\epsilon_s q N_A} \cdot (V_{gs} - V_{FB} - V_{bs}) \right\}^{\frac{1}{2}} \right] \quad (16)$$

where

$$\Phi_f = -\frac{kT}{q} \ln \left(\frac{N_A}{n_i} \right)$$

Equation 16 better calculates V_{DSAT} and does not lose its validity when the inversion layer disappears near the drain, i.e., when $V_D = V_{DSAT}$.

It is important to accurately model two diodes (body-source and body-drain) because they significantly impact subthreshold behavior. The I-V characteristics of two diodes are modeled as follows {15}.

$$I_{diode} = I_{diff} + I_{gen} \quad (17)$$

where

$$I_{diff} = q \cdot A_j \cdot \frac{n_i^2}{N_A} \sqrt{\frac{D_n}{\tau_d}} \cdot (e^{\beta V_d / n_d} - 1)$$

$$I_{gen} = q \cdot A_j \cdot \frac{n_i^2}{2\tau_g} \cdot W_{dep} \cdot (e^{\beta V_d / n_g} - 1)$$

where

- $A_j = W \times t_b$ - junction area
- τ_d - diffusion life time
- $V_d = V_{bs}$ for I_{BS}
- $V_d = V_{bd}$ for I_{BD}
- W_{dep} - width of depletion region
- τ_g - generation life time
- D_n - electron diffusion coefficient
- n_d, n_g - diode constants
- $n_d \approx 1, n_g \approx 2$

In modeling the parasitic BJT, the recombination component of the base current, I_{BB} is considered because of the wide base width, L . Assuming a linear distribution of excess minority carriers in the base region, I_{BB} is given by [31]

$$I_{BB} \equiv I_B = q \cdot A_j \cdot \frac{n_i^2}{N_A} \cdot \frac{1}{\tau_b} e^{qV_{bs}/kT} \quad (19)$$

where τ_b is the recombination lifetime of minority carriers in floating body acting as a base. The emitter and collector currents are

$$\begin{aligned} I_E &= (h_{FE} + 1)I_B \\ I_{CE} &= h_{FE} \cdot I_E \end{aligned} \quad (20)$$

where h_{FE} is the dc current gain.

Thus, we obtain from Fig. 4, the drain-to-source, SOI n-channel MOSFET current

$$I_{DS} = I_D + I_{BS} + I_E \quad (21)$$

Results and Discussion

Table 1 summarizes device and dose parameters of a typical n-channel SOI MOSFET. Calculations for I-V characteristics are divided in two parts. For a given V_{gs} , the first part calculates surface potentials and the second part finally calculates different current components: I_D , I_{IC} , I_{BS} , I_{BD} and I_{DS} for a given device, dose parameter values (A , D) and varying substrate bias. Effect of parasitic BJT can be included or ignored through a choice of h_{FE} . The programs are written in FORTRAN77 in UNIX operating environment. In modeling subthreshold behavior under total dose conditions, an average value of substrate bias, V_{bs} was considered. However, V_{bs} can also be obtained directly from eq.17 for I_{BS} .

Figure 5 summarizes the subthreshold I_D - V_{GS} characteristics for $V_D = 5$ V, $D = 0$ and $D = 10^5$ rad (SiO_2). Solid lines correspond to $D = 0$ and $V_{bs} = 0, 0.2$ and 0.5 V. Dotted lines correspond to $D = 10^5$ rad (SiO_2) and $V_{bs} = 0, 0.2$ and 0.5 V. It can be seen from the Fig. 5 that subthreshold behavior for $D = 10^5$ rad (SiO_2) changes significantly from its characteristics for $D = 0$ (pre-irradiated condition). The transistor enters into the strong inversion region from the weak inversion region. It should be noted that $V_{bs} = 0$ characteristics correspond to a typical bulk n-MOSFET. The pattern of shift in subthreshold characteristics agree with the measured total dose

response reported in literature [21]. Figure 6 shows the variation of current components: I_D , I_{DS} , I_{IC} and I_{BD} with varying gate bias voltages for $V_D = 5$ V, $V_{bs} = 0.2$ V and $D = 10^5$ rad (SiO_2). It should be noted from the Fig. 6 that $I_{IC} \gg I_{BD}$ for the SOI n-MOSFET parameters.

Conclusions

An n-channel SOI MOSFET model is presented which explains the shift in its subthreshold I-V characteristics as a result of total dose radiation response. Radiation induced trapped positive charge and interface states charge models are included in modeling of MOSFET's based on charge-sheet model. The equivalent SOI n-MOSFET model includes both the body-source, body-drain diodes and the parasitic BJT. It is assumed that threshold voltage of back-gate parasitic MOS transistor is more than V_{SUB} . The model predicts the shift in subthreshold behavior with total dose and agrees with the observed behavior reported in literature.

The present n-channel SOI MOSFET model is continuous from subthreshold to strong inversion region. Thus the model can be used to predict characteristics from weak inversion to linear and saturation regions against the total dose using device and dose parameters as the input. The model can also be used to predict radiation hardness of SOI/CMOS devices and circuits using basic parameters.

Future work is planned to compute subthreshold behavior for other dose related parameter values (A , D) than the one used in the present analysis. The cases of $I_{IC} \gg I_{BD}$ and $I_{IC} \ll I_{BD}$ need to be examined, to determine the effect of various device parameters and varying drain bias conditions. Verification of the present analysis and modeling through SPICE simulation and experimental verification are proposed for the future work.

Acknowledgments

The investigator acknowledges the Air Force Office of Scientific Research for the summer 1996 research award under Summer Faculty Research Program. Dr. W. Shedd, Focal Point for the program is thankfully acknowledged for his support, interest, weekly discussions and suggestions for the proposed research. The investigator acknowledges Dr. B. K. Singaraju and Capt. S. Clark for their interest and suggestions, Drs. S. P. Karna, J. Garth, C. Brothers, Mr. J. Chavez and Mr. W. Kemp, Dr. N. Islam, Dr. R. Pugh and Prof. E. Sanchez-Sinencio of Texas A & M University for their interest. Last but not the least Ms. J. Griego and Ms. C. Duggins are thankfully acknowledged for their administrative support and Mr. F. Jaramillo for computer related services.

References

1. T. Nishimura, Y. Akasaka, and H. Nakata in *Silicon-on-Insulator: Its Technology and Applications*, S. Furukawa (Ed.), pp. 263-268, KTK Scientific Publishers, Tokyo, 1985.
2. P. K. Vasudev, *Solid State Technology*, pp. 61-65, Nov. 1990.
3. J. P. Collinge, *Silicon-on-Insulator Technology: Materials to VLSI*, Kluwer Academic Publishers, Boston, 1991.
4. R.A. Metzger, *Compound Semiconductor*, vol. 2, pp. 44-45, May/June 1996.
5. D. Neamen, W. Shedd and B. Buchanan, *IEEE Trans. Nucl. Sci.*, vol. NS-21, pp. 211-216, Dec. 1974.
6. M.C. Biwer, Total-Dose Response of Silicon-on-Insulator (SOI) Metal-Oxide-Semiconductor Transistor's (MOSFET's), M.S. Thesis, Electrical Engineering Department, Northeastern University, Boston, MA, June 1988.
7. N.K. Annamalai, W. Shedd, J. Chapski and T. Kearney, *1991 IEEE Int. SOI Conf. Proc.*, Oct. 1-3, pp 146-147, 1991.
8. P.S. Winokur, D.M. Fleetwood and F.W. Sexton, *Radiat. Phys. Chem.*, vol. 43, no. 1/2, pp. 175-190, 1994.
9. H.K. Lim and J.G. Fossum, *IEEE Trans. Electron Devices*, vol. ED-32, pp. 446-457, 1985.
10. S. P. Edwards, K.J. Yallup and K.M. De Meyer, *IEEE Trans. Electron Devices*, vol. 35, pp. 1012-1020, July 1988.
11. S. Veeraraghavan and J.G. Fossum, *IEEE Trans. Electron Devices*, vol. 35, pp. 1866-1875, Nov. 1988.
12. J.G. Fossum, *IEEE Trans. Electron Devices*, vol. 37, pp. 724-729, March 1990.
13. R. Howes, W. Redman-White, K.G. Nichols, M. Robinson and J. Kerr, *1991 IEEE Int. SOI Conf. Proc.*, pp. 90-91, Oct. 1-3, 1991.
14. M.B. Abel, M. Berger and G. Zimmer, *1992 IEEE Int. SOI Conf. Proc.*, pp. 42-43, Oct. 6-8, 1992.
15. M. Matloubian, C.E. D. Chen, B.Y. Mao, R. Sundaraesan and G.P. Pollak, *IEEE Trans. Electron Devices*, vol. 37, pp. 1985-1994, Sept. 1990.
16. A. Srivastava and K. Venkatapathy, *VLSI Design*, vol. 4, no. 1, pp. 75-81, 1996.
17. A. Srivastava and K. Venkatapathy, *Cryogenics*, vol. 35, no. 9, pp. 599-605, 1995.

18. T. Ohno, K. Izumi, M. Shimaya and N. Shiono, *IEEE Circuits and Devices Magazine*, vol. 3, pp. 21-25, Nov. 1987.
19. T. Ouisse, G. Ghibaudo, J. Brini, S. Cristoloveanu and G. Borel, *IEEE Trans. Nucl. Sci.*, vol. 39, pp. 372-375, June 1992.
20. J.H. Smith, R. Lawrence and G.J. Campisi, *Proc. of the Fifth Int. Symp. on Silicon-on-Insulator Technology and Devices*, pp. 171-177, May 1992.
21. T.P. Ma and P. Dressendorfer, *Ionization Radiation Effects in MOS Devices and Circuits*, John-Wiley, NY, 1989.
22. H.E. Boesch, Jr., F.B. McLean, J.M. Benedetto, J.M. McGarrity, and W.E. Bailey, *IEEE Trans. Nucl. Sci.*, vol. NS-33, pp. 1191-1197, Dec. 1986.
23. R. Freeman and A. Holmes-Siedle, *IEEE Trans. Nucl. Sci.*, vol. NS-25, pp. 1216-1225, Dec. 1978.
24. N. Shino, M. Shimaya and K. Sano, *Jap. J. Appl. Phys.*, vol. 22, pp. 1430-1435, Sept. 1983.
25. C.E. Barnes, D.M. Fleetwood, D.C. Shaw and P.S. Winokur, *IEEE Trans. Nucl. Sci.*, vol. 39, pp. 328-341, June 1992.
26. J.R. Brews, *Solid-State Electronics*, vol. 21, pp. 345-355, 1978.
27. S.M. Sze, *Physics of Semiconductor Devices*, John-Wiley, NY, 1969.
28. Y.A. El-Mansy and D.M. Caughey, *Int. Electron Device Meeting Tech. Digest* (Wash. D.C., 1975), pp. 31-34.
29. J.G. Fossum, J.Y. Choi and R. Sundaresan, *IEEE Trans. Electron Devices*, vol. 37, pp. 724-729, March 1990.
30. A. S. Grove, *Physics and Technology of Semiconductor Devices*, John-Wiley, NY, 1967.
31. D.A. Hodges and H.G. Jackson, *Analysis and Design of Digital Integrated Circuits*, McGraw-Hill, NY, 1988.
32. P. Villard and R. Kielbasa, *1996 IEEE Proc. of RADECS '95*, pp. 33-39.
33. M.D. Jacunski and M.C. Peckerar, *IEEE Trans. Nucl. Sci.*, vol. 39, pp. 1947-1952, Dec. 1992.
34. D. Kahngi (Ed.), *Silicon Integrated Circuits - Part A, Applied Solid State Science, Supplement 2*, Academic Press, NY, 1981.

Table 1: SoI n-MOSFET device parameters

t_{ox}	$= 20 \text{ nm}$
x_1	$= 5 \text{ nm}$
N_A	$= 1 \times 10^{17} \text{ cm}^{-3}$
N_D	$= 1 \times 10^{20} \text{ cm}^{-3}$
W	$= 3.6 \text{ } \mu\text{m}$
L	$= 1.2 \text{ } \mu\text{m}$
μ	$= 300 \text{ cm}^2 / \text{v} - \text{s}$ (for poly-Si/n ⁺ - doped)
t_b	$= 140 \text{ nm}$
τ_d	$= 1 \text{ ns}$
τ_g	$= 1 \text{ ns}$
τ_b	$= 1 \text{ ns}$
β	$= 2$
Q_{ss} / q	$= 2 \times 10^{10} \text{ cm}^{-2}$

Circuit parameters:

V_{ds}	$= 5V = V_D$
V_S	$= 0$
V_G	$= 0 \text{ to } 1V$

Dose parameters:

A	$= 1$
D	$= 1 \times 10^5 \text{ rad}(\text{SiO}_2)$

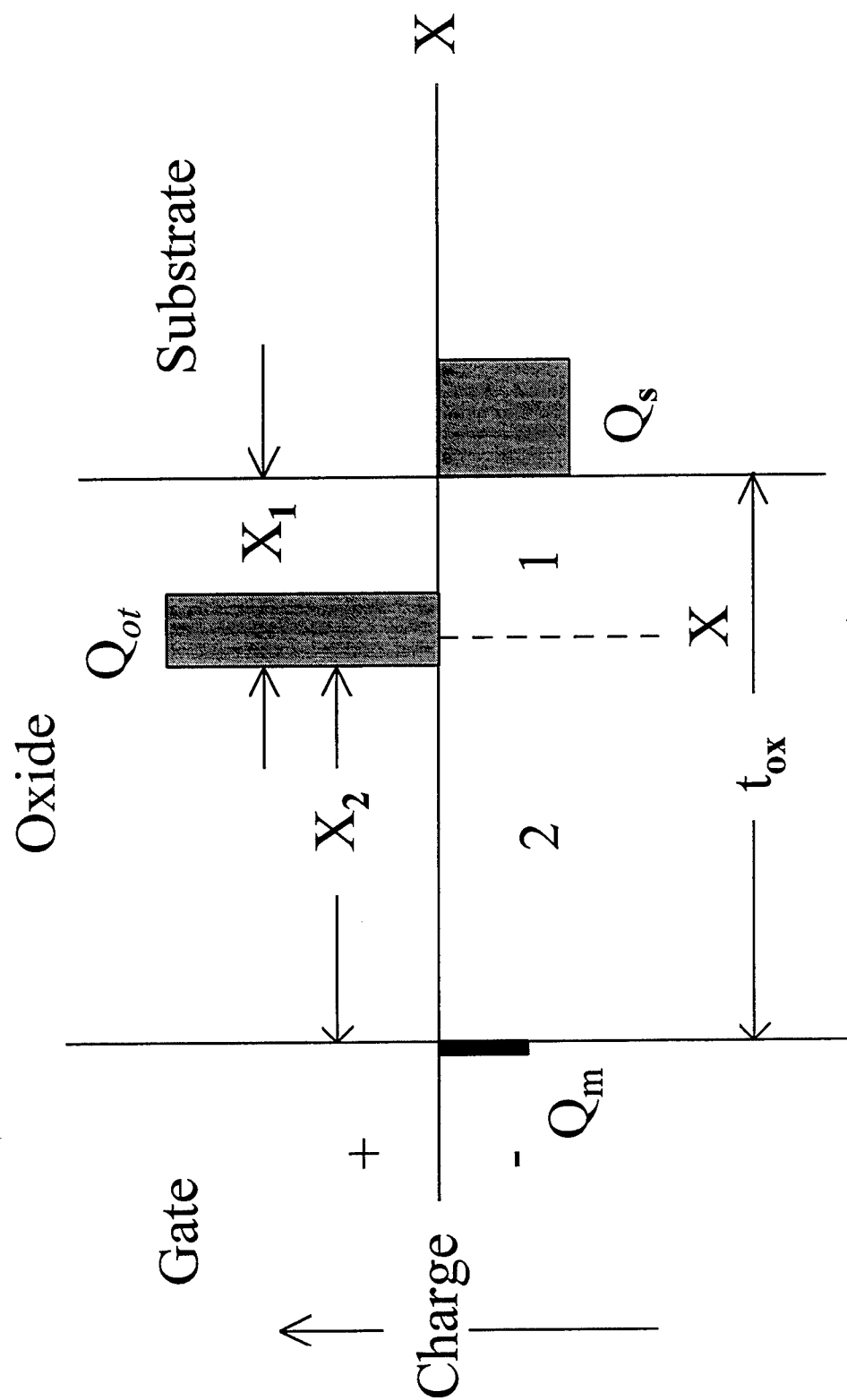


Fig. 2. A mathematical model for radiation-induced oxide charge

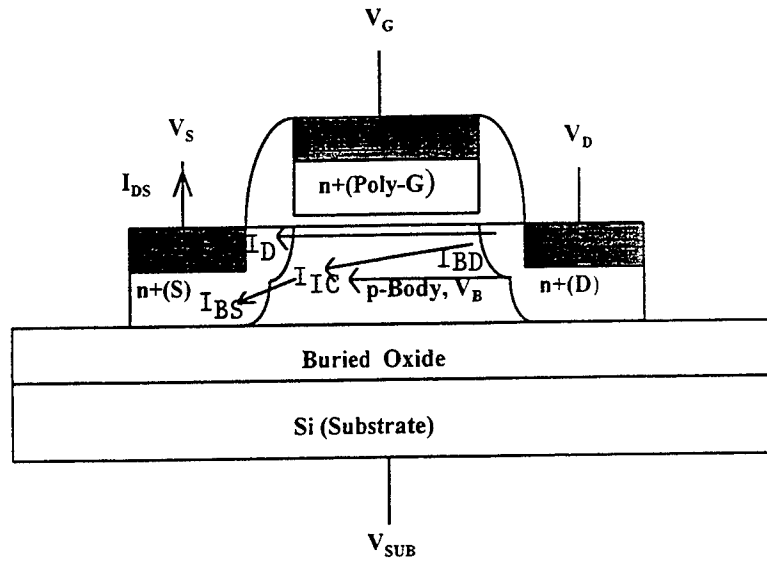


Fig. 3. Cross-section of an n-channel SOI MOSFET

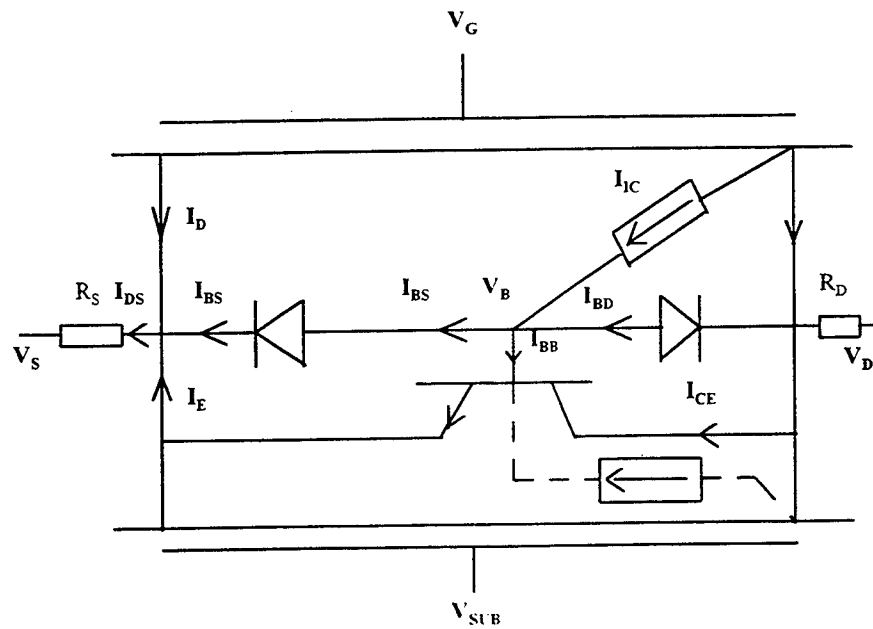


Fig. 4. Equivalent circuit model of an n-channel SOI MOSFET

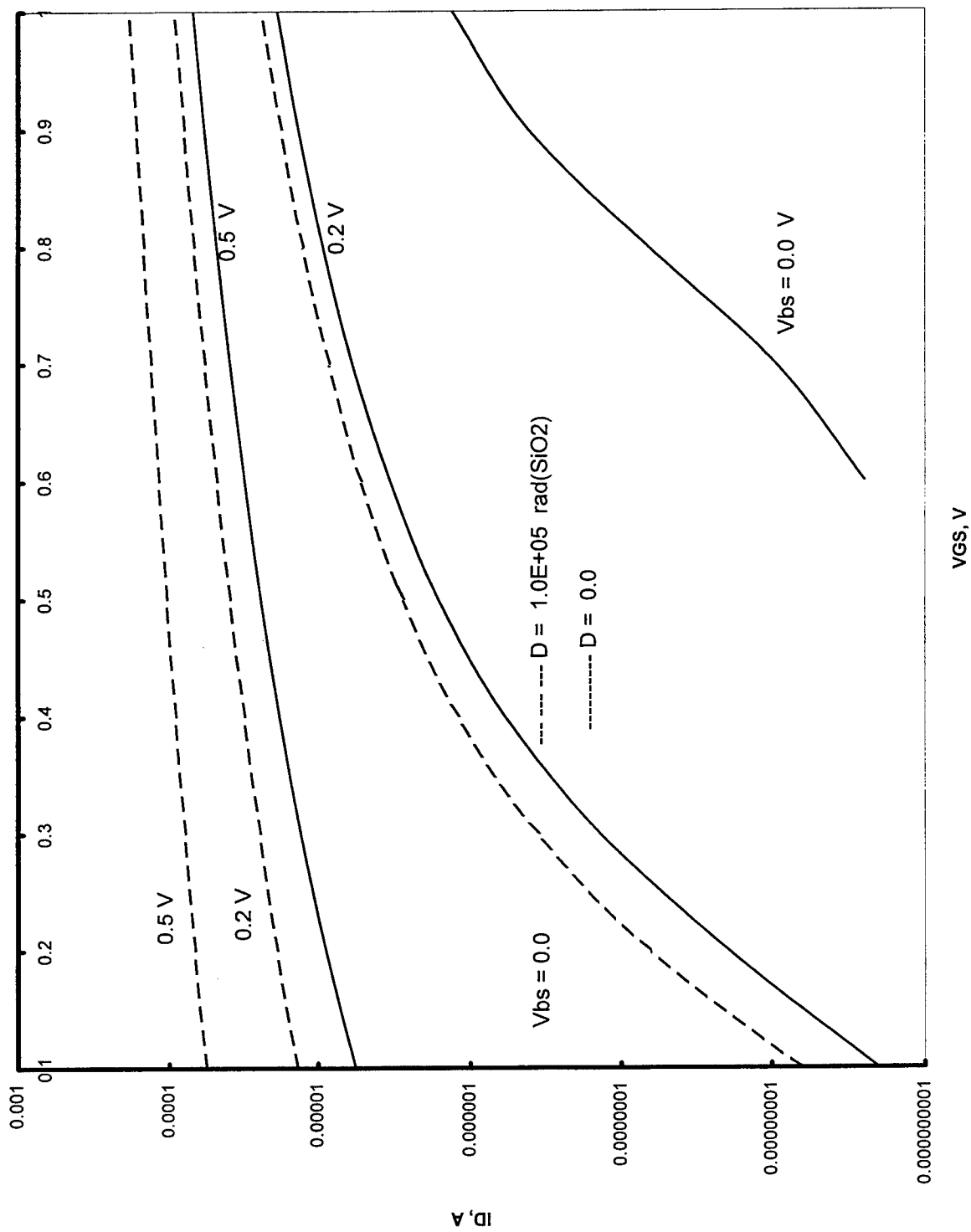


Fig. 5. Subthreshold $I_D - V_{GS}$ characteristics with varying V_{bs}

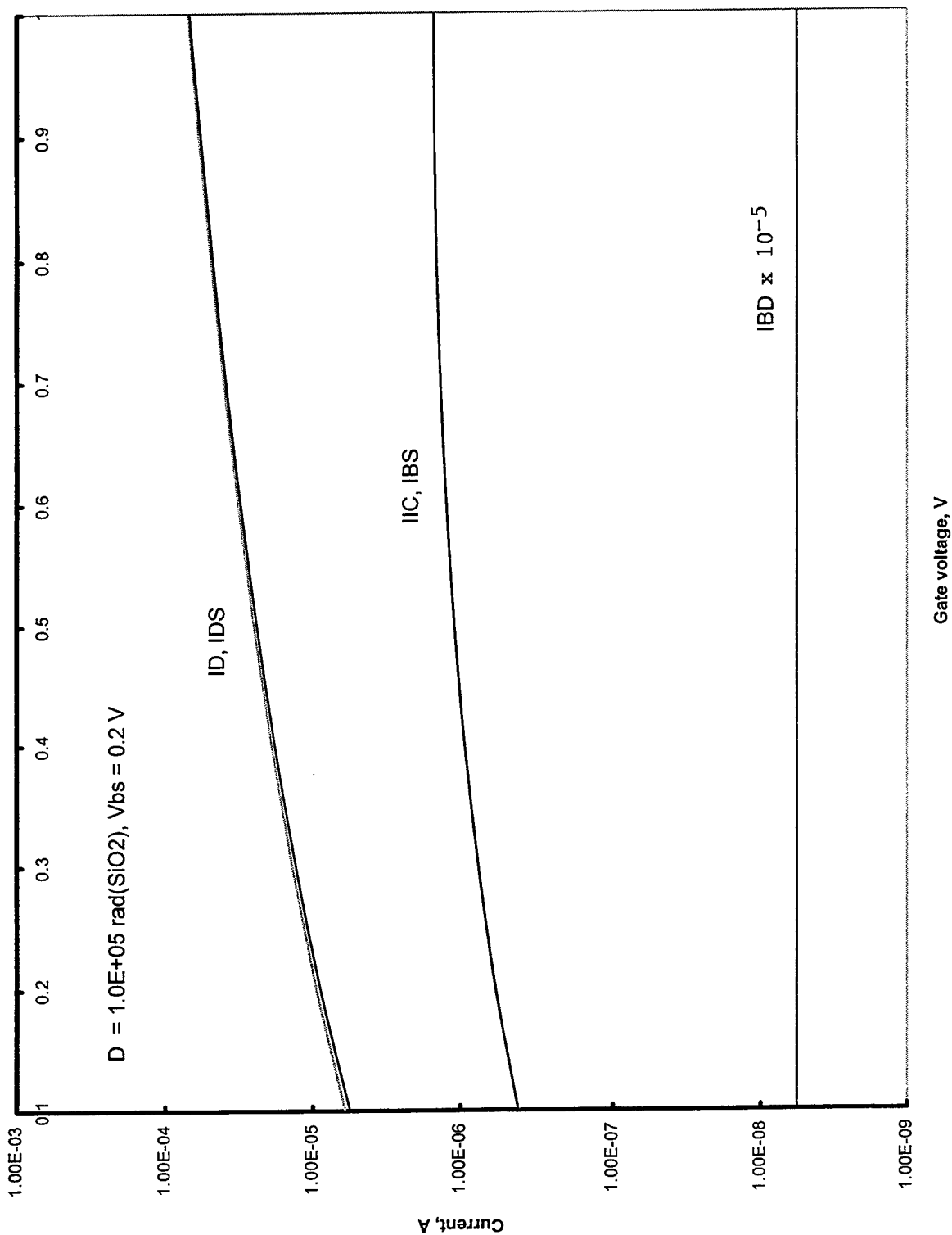


Fig. 6. Variation of Current components with V_{GS}

**THE POTENTIAL APPLICATIONS OF SUPER-RESOLUTION
AND ARRAY PROCESSING TO SPACE-BASED RADARS**

**James M. Stiles
Assistant Professor
Electrical Engineering and Computer Science
The University of Kansas**

**2291 Irving Hill Drive
Room 310
Lawrence, Kansas
66045**

**Final Report for:
Summer Faculty Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC**

September 1996

THE POTENTIAL APPLICATIONS OF SUPER-RESOLUTION AND ARRAY PROCESSING TO SPACE-BASED RADARS

James M. Stiles
Assistant Professor
Electrical Engineering and Computer Science
The University of Kansas

Abstract

To justify placing a radar in space, the sensor must exhibit high performance over a number of radar modes, yet provide these goals with a sensor small and light enough to be economically placed in space. To achieve this, a sensor with increased hardware complexity and computational power is likely required. By designing the receive aperture as an instrument for collecting both power and information, the potential exists to collect sufficient information to provide high-resolution detection of targets, whether stationary and moving, airborne or on the earth's surface. Additionally, larger area search rates than provided by conventional radars can be obtained. This is achieved by implementing multiple apertures and sub-apertures as the receive antenna, the output of each independently and coherently sampled. This volume of data provides an extreme amount of information, and can be used to implement such functions as digital beamforming and super-resolution.

1 INTRODUCTION

The many benefits of placing active microwave sensors in space are apparent. A radar can provide surveillance over a huge area, and in all weather conditions. The sensor does not require a crew, nor does it require refueling. It can operate continuously, observing all portions of the globe. Yet, relatively few space-borne radars have been implemented, the preponderance of these sensors are either ground-based or fly on aircraft platforms (AWACS, J-STARS, etc.). The reasons for this dearth of space-borne sensors are due to the many technical challenges associated with this task. The normal problems associated with space, i.e. power and weight, are especially difficult for a space-borne radar (SBR). In contrast to passive optical and infrared systems, the active SBR must generate and transmit sufficient energy to detect targets, a particular challenge at the extreme distances associated with space applications. Additionally, the relatively long wavelengths of the microwave region result in comparable apertures which are physically very large. Not only does this increase sensor weight, but additionally makes deployment of these large structures problematic.

The problems of weight, size, and power are unavoidable when considering SBR. The question becomes then, how to build a sensor with a specified performance, such that these parameters are minimized. As with any engineering problem, the SBR designer will be faced with performance and cost trades. The challenge is to trade weight, size and power with other sensor parameters while preserving sensor performance. Two obvious candidates for these trade-offs are processing power and hardware complexity. The implementation of technologies such as MMICs have lead to sensors with additional capabilities at a reasonable increase in cost (e.g. active phased-arrays with MMIC T/R modules).

Additionally, the continuing increase in computing performance and speed is well known. Perhaps a combination of these technologies could be utilized in the future to provide acceptable SBR performance with reasonable power and size requirements. The result could be a sensor whose operation is quite different than developed for ground-based or airborne radars, sensors whose performance and cost criteria are not as directly affected by the parameters of size, weight, and power. Thus, it would seem important when considering SBR, to not proceed only with the traditional methods of radar design and processing. Otherwise, the feasibility of spaced-based radar could perhaps be incorrectly evaluated, as implicit in the conclusion would be a dependence on standard methods and designs.

This paper will propose SBR design methods which could potentially provide high sensor performance over a number of operating modes, while requiring reasonable power, size, and weight specifications. The proposed methods, however, do place an increased burden on processing requirements and hardware complexity. The assumption is that this trade will provide a positive impact on both the cost and feasibility of spaced-based radar systems.

2 SENSOR REQUIREMENTS

When comparing the relative performance of radar sensors, it is important to note the requirements and assumptions which were used to synthesize the design. Table 1 provides this information for this paper, a table which demonstrates the extreme requirements faced by a SBR designer. One such requirement is the area search rate. Assume a theater search area of approximately 2.6×10^5 square kilometers, and that the SBR must provide an new update of this search area every 10 seconds. The result is a huge search area rate

Target rcs	2	m ²
Search Area	2.6×10^5	km ²
Update Rate	10	sec
Area Search Rate	2.6×10^4	km ² /sec
Resolution	15	meters
Min. Discernible Vel.	2	m/sec (radial)
Sensor Altitude	1600	km
Radar Velocity	7000	m/sec

Table 1: *The numeric requirements for SBR used in this study. Although the exact accuracy of these values are certainly debatable, this set of requirements hopefully is representative of the technical challenges associated with SBR.*

of 2.6×10^4 square kilometers/second. A second challenge is to detect a target with an rcs of 2 square meters at a slant range of 3200 kilometers. Finally, the radar must detect moving targets with a radial velocity as low as 2 meters/sec.

To provide a basis for this analysis, some radar performance and scenario parameters were likewise assumed. These are listed in Table 2, and were conservatively selected to insure the feasibility of the resulting design. A possible exception to this is the average power specification of 10 kW. Although current technologies would be unable to deliver this power in space (at least with reasonable weight and cost), it is assumed that the projected advances in power technology will lead to space systems which can deliver this requirement. Note that 10 kW, although very large, is at the low end of what is often

Frequency	10	GHz
Average Power	10	kW
Noise Figure	5	dB
Grazing Angle	30	degrees
Required SNR	18	dB
Slant Range	3200	km

Table 2: *The assumptions (both radar and scenario) used in this study. The key assumption presented is the average power parameter.*

perceived as necessary to satisfy the SBR requirement. This value, along with the other parameters presented in these tables, are certainly debatable. The point of this paper is not to present these numbers as necessary and sufficient for an optimal SBR design. Rather, the data presented in this section was selected to provide a reasonable parameter set from which to demonstrate the potential of the ideas and methods presented later.

3 SBR DESIGN

This section presents a cursory design for a spaced-based radar to meet the requirements of Table 1 using the assumptions of Table 2. Although obviously not a rigorous or detailed design, it is intended to determine the potential of specific radar systems and algorithms for satisfying the design requirements with a relatively small and low power radar.

The design requirements will be interpreted in the broadest sense: to detect targets

and determine with precision their location and velocity, whether they be stationary or moving, airborne or ground targets. Note that this interpretation covers a range of traditional radar operating modes. It satisfies the both the GMTI and AMTI mission, while also providing surface imaging and mapping data (natural scatterers are simply stationary ground targets). By stating the radar requirement in this broad manner, it avoids the temptation to link radar function (imaging, GMTI, etc.) to a specific method or implementation (SAR, DPCA, etc.). It could be argued that the requirements perceived necessary to satisfy a given radar function are often an artifact of the *methods* assumed to perform that function, and not of the inherent physics/mathematics of the function itself. In order to satisfy the requirement previously described, the proposed design implements hardware and algorithms which cannot in general be assigned to a specific methodology, but rather uses and extends several known methods in perhaps unorthodox ways.

For a monostatic radar design, the receiver and transmitter typically share an antenna aperture, thus the antenna is designed to maximize overall system performance and utility. However, the result may be a compromise between the antenna performance as a receive aperture, and its performance as a transmit aperture. On transmit, an antenna acts essentially as an energy focusing device, designed to place all available power onto the targets of interest, while illuminating no others. It could be argued that the receive antenna has a similar function, to efficiently receive energy from the targets of interest and no others. This is certainly one of the functions of the receive antenna, however in many cases the receiver is tasked to do more.

A transmitter will emit a known signal, where the propagating field can be described *a priori* at every point in time and space. Conversely, the scattered fields are largely unknown, as the known incident fields are modified by the unknown parameters of the

illuminated targets. A receive antenna must be constructed which therefore sufficiently senses the returning scattered field, such that a receiver can infer the requisite target parameters from the signal supplied from the receive antenna. As the number of required target parameters to be inferred increases, more information must be supplied by the receive aperture to insure high accuracy and low ambiguity of these target parameter estimates. This requires that the receive antenna sense the scattered electric field over a wider region of both space and time.

Therefore, whereas the transmit antenna is an energy focusing device, the receiver antenna can be described as an information collection device. Collecting sufficient energy such that receiver noise does not obscure this information is indeed important, but is far from the only function of the receiver aperture. Given that the transmit and receive aperture provide differing functions, it is perhaps logical to conclude that using the same antenna for both transmit and receive is a constraint which can limit overall radar performance. Therefore this design will assume that these apertures are both separate and dissimilar.

The size of the radar antenna (both transmit and receive) is among the most important parameters of a SBR design, as it usually is considered to be the structure which drives total system weight. Additionally, the size of the antenna will most certainly be larger than the dimensions of the launch vehicle payload, thus requiring a mechanical deployment in space. The difficulty of this challenge is at the very least proportional to the physical size of the aperture. Some proposed SBR designs use very large aperture sizes to satisfy the system requirements, with diameters approaching 100 meters. The great size of these proposed systems are due to the three functions which the radar aperture provides: 1) it focuses the transmitted power on the target (provides gain), 2) it collects the

scattered power from a target (effective aperture), and 3) it provides spatial resolution, illuminating and receiving power from a small spatial area. Of course, for all three of these functions, performance increases with aperture size. Conversely, these functions are adversely impacted as slant range increases, therefore at the extreme ranges associated with space radar, the required antenna size must increase to compensate.

However, the difficulty in SBR implementation arises from the *physical* size of the antenna. This physical size need not limit the aperture size of the radar, for the concept of a synthetic aperture is well known and often implemented. For the assumed altitude of this design, an array 7000 meters wide can be synthesized in just 1 second. Of course, the physical size of the antenna is still important, as are the individual elements in any array, however the requirements on the physical antenna are changed and perhaps diminished if a synthetic aperture is assumed.

Synthetic aperture radars are most often associated with surface imaging, and thus not often considered when designing for other radar functions, such as MTI. However, a synthetic aperture can be used to focus on a moving (wrt the ground) target as well, providing both resolution and good SNR. Unfortunately, the response of a synthetic aperture will likely exhibit ambiguity between the position and the velocity of a moving target. A SAR sensor can equivalently be described as a pulse-doppler radar, with a doppler resolution equivalent to the resolution obtained with the synthetic aperture. As the observed doppler is a function of the relative radial velocity between target and radar, the doppler associated with a moving target is a function of both its position and velocity. Thus, with a single synthetic aperture observation, it appears that a moving target could be largely indistinguishable from a stationary target at a different location. Indeed, because of the down-looking geometry of the space-based problem, the scattered energy from a moving

target will almost surely share the same resolution cell as a stationary clutter targets. Although focusing the aperture on a moving target will result in some attenuation of the received signal from stationary targets, it is likely that the target to clutter ratio will still be insufficient. Unless a method can be found for resolving the two responses, it would seem that using a synthetic aperture would preclude the detection of moving targets. However, the benefits of synthesizing a very large synthetic aperture would suggest that such a method be sought.

3.1 Super-Resolution

By synthesizing a very large aperture (or equivalently constructing a range-doppler map), a SBR provides a high-resolution measurement cell indicating scattered energy from one and perhaps two sources. The first, which will almost always be present, is the scattering from a section of the earth's surface defined by the ambiguity function of the observation (i.e. clutter). The second, which may or may not be present, is the scattering from moving (wrt the ground) targets. If a moving target (with a non-zero radial velocity wrt the radar) is present, it is apparent that it cannot be colocated with the clutter target; a ground moving target will always generate a different doppler than the surface underneath it. The scattered energy from each target will therefore arrive at the radar from slightly different incidence angles, as the energy from the two scatterers originate from separate locations on the equi-range contour.

Angle-of-arrival, therefore, can potentially be used to resolve the two responses (it is assumed that at most there will be one moving target in each high-resolution cell). Generally, to resolve scatterers with dissimilar incidence angles, a receive antenna with a narrow beamwidth is implemented such that the incident power from one target is

efficiently coupled into the receiver, while rejecting the power of the other. The problem is that this aperture cannot be synthesized; it is by synthesizing an aperture over time that the original ambiguity is generated. Instead, the physical aperture must be large (i.e. narrow beamwidth) enough such that only the scattered power from the moving target is coupled into the receiver as the aperture is synthesized. In addition to requiring precise beam steering to continuously point at the moving target (and away from the stationary clutter), such a solution would counter the original design goal - to reduce the physical aperture size. For ground moving targets with small radial velocities, the distance between the moving target and the clutter target which shares the same resolution cell is small (they are obviously colocated if the velocity is zero). Therefore, to spatially resolve the two targets would again require a very large physical aperture.

However, it is possible to provide angular resolution far superior to what an aperture size would suggest. Two examples of this are monopulse radars and phase interferometers. These are but two methods which can be generally referred to as super-resolution. The direction-of-arrival of an incident wave is determined from a set of observations collected from dissimilar antennas. These dissimilar observations are the key to super-resolution, and why angle-of-arrival estimates are capable of significantly greater accuracy than the sensor size alone would suggest. For example, consider an array of antenna elements over some circular area of diameter d . By coherently adding the output of each array element, an antenna is constructed which has a beamwidth proportional to λ/d . The angular resolution of a sensor using this antenna is therefore just the beamwidth of the antenna; if a signal is detected, it is assumed to be within the beamwidth. If however, the output of each individual element is coherently detected, more information is available in which to infer the angle-of-arrival of the incident wave. This information, such as the relative phase

differences between each element, can be used to infer the incident direction of the wave. However, it is destroyed if coherently added to provide a single antenna port, destruction which reduces angular measurement accuracy to the beamwidth of the antenna.

Although monopulse and interferometry techniques are generally constructed assuming a single incident signal, techniques such as the Multiple Signal Classification (MUSIC) algorithm (among others), have been derived which allow for the resolution of two (or more) incident signals. These techniques, of course, require multiple dissimilar observations (multiple receivers) and additional processing power. However, the trade between hardware/processing complexity and aperture size is likely a desirable one for SBR. To meet the design requirements described in Section 2 (minimum discernible velocity), at least five receiver elements are required, with the two most distant elements being on the order of 20 meters apart.

An example of the such a sensor is sketched in Figure 1, exhibiting a design wherein five receiver elements are implemented. Five dissimilar receive apertures can thus be synthesized, with each providing a different complex response. From these ten real values, a moving target can potentially be detected, and its location inferred, along with the location and scattering response of the clutter target. Note that the location vector can be provided not just in azimuth, but in elevation as well. This is an important value when considering airborne moving targets, as a large uncertainty in target location would otherwise result. If only azimuth and range to target were provided by the radar, then an uncertainty would remain in terms of elevation angle.

This uncertainty is bounded on one end by the earth's surface, and on the other end by the maximum altitude of aircraft. However, this uncertainty region is not ver-

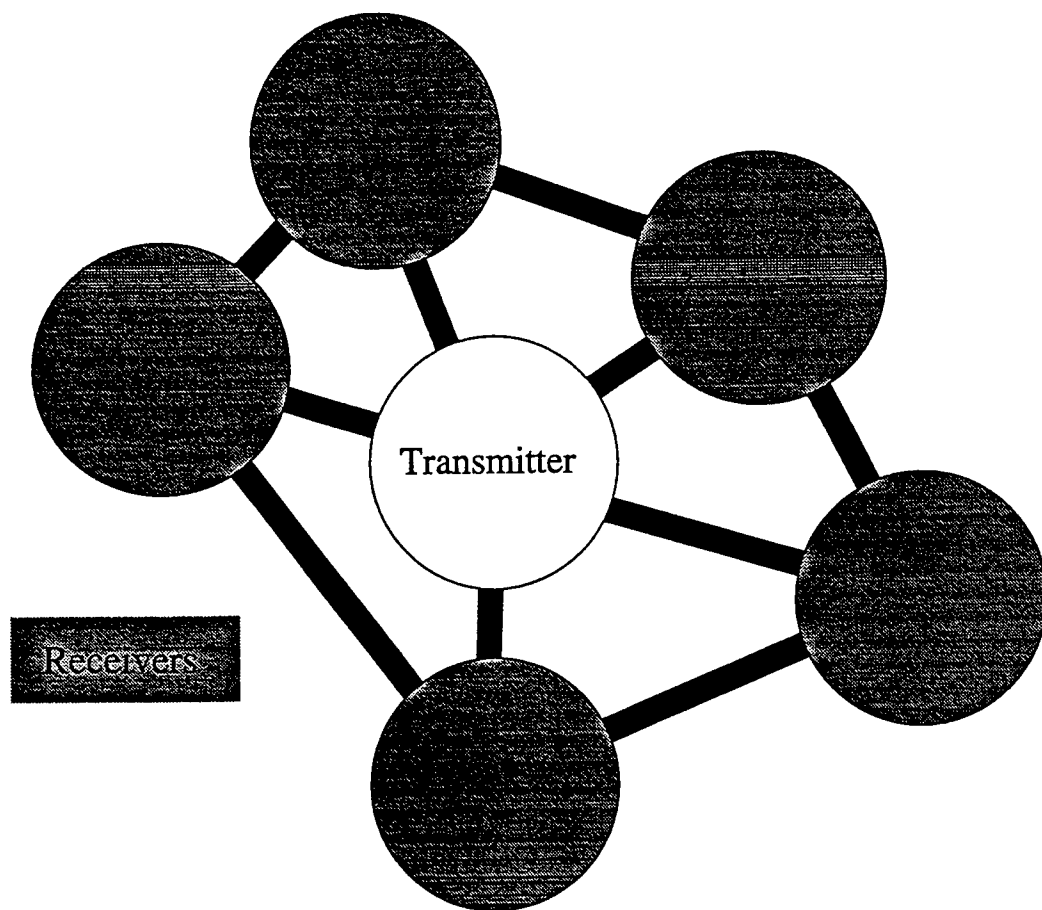


Figure 1: *Possible configuration of an SBR sensor, showing multiple receive aperture for super-resolution processing.*

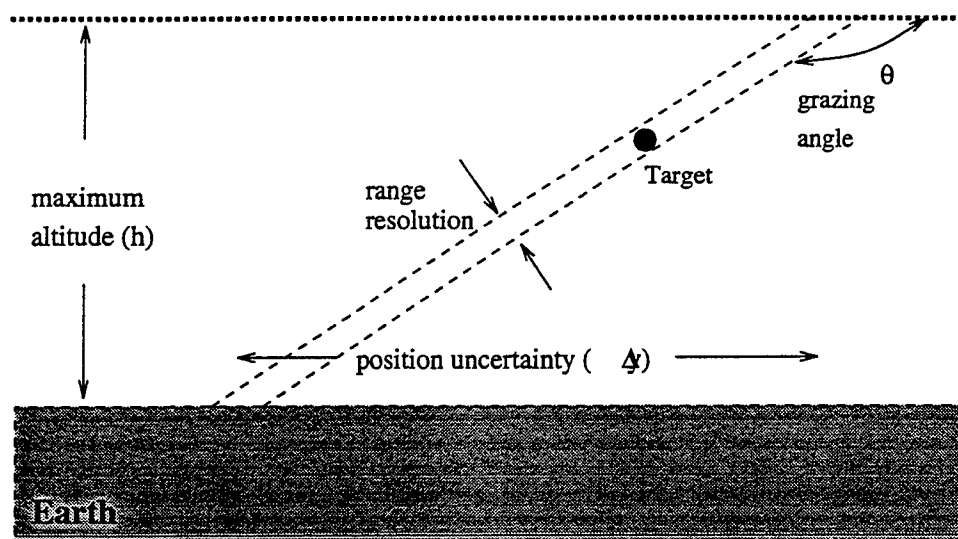


Figure 2: Geometry showing the uncertainty in position resulting from no elevation measurements of an airborne target.

tical with respect to the earth, but tilted by the grazing angle of the incident wave. As shown in Figure 2, when projecting this column onto the earth's surface, one finds that the uncertainty in elevation angle translates into a uncertainty in ground position of $\Delta y \approx h_{atmosphere} \tan \theta_{grazing}$, where $h_{atmosphere}$ is the height of the navigable atmosphere, and $\theta_{grazing}$ is the radar grazing angle. Assuming $h_{atmosphere} = 18$ kilometers and $\theta_{grazing} = 45$ degrees, this uncertainty would be 18 kilometers, even for radars with small range resolution. Therefore, the benefit of providing good range resolution is significantly diminished, as resolution cannot be translated to spatial accuracy unless an elevation estimate (providing the third spatial dimension) is determined. Assuming a maximum 25 meter separation of the receive antennas, the uncertainty in ground position could potentially be reduced to a 350 meters ($\theta_{grazing} = 45^\circ$).

3.2 Transmit Antenna

The concept displayed in Figure 1 shows a single transmit antenna whose specifications have not yet been discussed. Whereas it is true that a large aperture implies large gain, such that the power density incident on a given scatterer is proportional to aperture area, this increase is at the expense of illumination spot size. Thus, for a given area search rate, the illumination time on a specific target is inversely proportional to the aperture size. Given that the relevant parameter for target detection is proportional to the product of power density and illumination time (i.e. energy), the effects of aperture size cancel, and the area of the transmit aperture is apparently irrelevant to sensor performance (again, assuming a constant area search rate). Therefore, the same SNR can be accomplished using a large, high-gain aperture and a small illumination time, or a larger illumination time coupled with a lower gain (smaller) antenna.

Of course, there are other issues which effect transmit aperture size. The search area assigned to a radar may be peculiarly shaped or even discontinuous. The transmit aperture must be large enough (i.e. narrow beamwidth) such that only this area of interest is illuminated. A small transmit antenna may result in a situation wherein the area of interest cannot be scanned without illuminating significant areas outside the assigned boundary. This illuminated energy is thus wasted, and either the received SNR or effective area search rate will suffer.

This is particularly true if the designated search area is smaller than the designated track area. Some SBR operational concepts include a situation where a search area is established around the border of a larger area. In this search area inbound targets are initially detected and tracked, but are continued to be tracked after they leave the search

area. Thus, the area of interest increases, as the radar must illuminate not only the entire search area but additionally the small area around each tracked target. Again, the radar should efficiently illuminate this added area, a requirement leading to a minimum aperture size in this case. The exact value of this minimum aperture is dependent on the number and distribution of the tracked targets in this scenario.

3.3 Receiver Aperture

Of the three reasons for constructing a large physical aperture, only one remains. Resolution is provided by constructing a synthetic aperture. The received energy (i.e. SNR) from a single target, for a given time and search area rate, is independent of transmit aperture size, provided an efficient scan can be accomplished. However the third reason, to collect and couple scattering energy into the receiver, is a function which is directly proportional to the effective aperture size of the receive antenna. It is this requirement which therefore drives the size of the antenna aperture. Under the conditions and assumptions of this design scenario, a receive aperture of approximately 65 square meters (9 meter diameter) is required.

Viewing the receive antenna as an information collection device is underscored when considering the high area search rates and fine resolution specifications which are desired for SBR sensors. The combination of these two requirements leads to a radar receiver with a very high output data rate, which can be expressed in terms of resolution cells (pixels) per second. For example, a 15×15 meter resolution cell would result in a required data rate of 115×10^6 samples per second. A data rate this large would at first seem unachievable, as rate is set by the bandwidth of the radar, in this case 10 MHz, or 10×10^6 samples per second. Attempts to use this data stream to provide a greater search area rate would

result additional but dependent data, and thus the inferred results would be wrought with ambiguity.

Note that the factor limiting the collection of information is the receiver. A transmitter can certainly be built which can illuminate the surface at the requisite search rate, a small transmit aperture, for example, could illuminate an arbitrarily large area. Again assuming that the transmit and receive antennas are separate and decoupled, this transmit antenna can likewise be scanned, thus illuminating a great extent of surface area in a given time. Since the transmit antenna is independent of its receiver, it could conceivably be transmitting on almost a constant basis. For example, it could transmit a pulse, immediately scan to another location on the surface and transmit again, repeating the process over and over. This of course would result in a tremendous increase in the scattered information incident on the receive aperture. Scattered returns from a great amount of targets would be arriving at the receiver, overlapped in time and incident from many directions.

However, a single receiver and aperture would generally be unable to collect all the information represented in these scattered fields. As stated before, a fundamental limit is the data sample rate of bandwidth B . Additionally, the receiver antenna efficiently couples incident energy from certain angular directions, as determined by the illumination pattern of the antenna. Thus, the scattered energy emitted from targets outside the mainbeam of the antenna will not be coupled into the receiver, thus further reducing the available information.

Consider instead a sensor with not just one receiver and aperture, but a collection of apertures, each with a separate receiver. The maximum data rate now available to

the sensor has increased from bandwidth B to NB , where N is the number of receivers in this collection. This increased data rate, however, does not insure an increase in the amount of available information. The additional data reflected in this increased data rate must be independent of the other data (sampling the same signal N times does not result in increased information). If however, the illumination pattern of each aperture is sufficient dissimilar, the receivers will sample scattering returns from different target locations. Thus, the increased sample rate could be used to provide a sensor of increased area search rate.

However, it would seem that this solution would likewise increase the overall aperture size of the sensor, the very parameter we are trying hard to reduce. This need not be the case, as this solution can be implemented using digital beamforming techniques. The outputs of an antenna array can be coherently weighted and added to provide a beam pattern in an arbitrary direction. By changing the complex weighting coefficients, other beam patterns can be produced. Different patterns can be generated at the same time by implementing parallel processors to coherently add the element outputs, essentially these processors are acting as spatial filters. Perhaps therefore, the easiest way to generate parallel receive patterns is to coherently detect each element output, then perform the beamforming in software. In this manner, multiple independent beam patterns can be generated (up to the number of elements used) without increasing the physical size of the aperture. To provide the area search rate given in Table 1, at least 12 independent beams (requiring ≥ 12 elements) must be provided on each of the receive aperture depicted by Figure 1.

Essentially, a sensor of this type could be viewed as several separate radars, with multiple transmitters sharing the same aperture (which is then time multiplexed) and likewise

with the receivers (which are spatially multiplexed). These multiple radars can thus increase the area search rate, but by how much is still uncertain. As with any radar, the problem which still must be overcome is ambiguity. Although never eliminated completely, the effect of scattering responses from ambiguous targets is minimized by matching the signal with the antenna illumination pattern. For example, the range and doppler ambiguities associated with a pulse-doppler radar are minimized by adjusting the prf until the ambiguous scatters lie outside the main beam of the radar antenna. Although not completely eliminated, the scattered energy from the offending targets are thus attenuated by the illumination pattern (an unambiguous prf cannot be found if the illuminated area is too large for the sensor bandwidth). A similar problem is encountered by the sensor described above, with the problem a function of antenna and scan patterns (both receive and transmit) as well as the radar signal. The problem is therefore more difficult, and may require new, perhaps arbitrary transmit signals to reduce target ambiguities to their minimum levels.

3.4 Conclusions

This report proposes a sensor which potentially could provide the high performance required of a space-borne radar, while achieving this with a minimum of total aperture size. The design presented here has a total (transmit and receivers) of approximately 400 square meters, yet has the potential to satisfy the requirements of Table 1. Potential modes include AMTI, GMTI and surface imaging. This performance relies on an increase in hardware and processor complexity, with a large amount of receive hardware implemented, and receive data collected, in order to perform the super-resolution and beamforming tasks. This increase in the total amount of received information, however,

can then be used to provide the data required by the SBR sensor. In other words, the receive aperture is used as an information (in addition to a power) collection device.

Note the adjective potential has been used here, the design presented here is merely a sketch, with values derived using approximate formulations. It is the purpose of this report, therefore, to present the potential of these ideas as relating to SBR. Much work is necessary to fully extend these ideas and provide a rigid theoretical framework. Once successfully completed, these ideas can be used to provide a rigorous design example.

Balloon Launch Retromodulator Experiment

Charles M. Swenson
Assistant Professor
Utah State University

Clark Steed
Graduate Student
Utah State University

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored By:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

October 22, 1996

Balloon Launch Retromodulator Experiment

Charles M. Swenson
Assistant Professor
Utah State University

Clarke Steed
Graduate Student
Utah State University

ABSTRACT

Under the AFOSR summer research program researchers from Utah State University / Space Dynamics Lab spent 12 weeks at the Phillips lab preparing and flying a retromodulator laser communication package on a high altitude balloon. The package was a prototype system for a low-power laser communications system for small low earth orbiting satellites. The work was divided into preparation of the ground station at Starfire Optical Range, assembly and testing of the retromodulator control electronics, micro controller programming and the actual balloon flight of September 15, 1996. All technical objectives of the experiment were met during the 1 1/2 hour flight of the balloon which reached a float altitude of 103,000 ft. The ferroelectric liquid crystal based retromodulator design of Utah State provided test patterns for modulation rates up to 20 kilo bits. Data was successfully down linked using a 1200 bps RS232 format and a simplistic receiver. This report outlines some of the reflected laser communications activities conducted under this summer research program.

Reflected Laser Communication Systems

Charles M. Swenson
Assistant Professor
Utah State University

Clark Steed
Graduate Student
Utah State University

1 Introduction

Satellite laser communications concepts have been under development for many years. Laser communications from ground to low earth orbit is relatively simple because most of the hardware is on the ground and the satellite need only be provided with a relatively simple sensor. However, conventional approaches to laser communication from satellite to ground require sophisticated hardware aboard the satellite such as a reliable laser subsystem and a capable subsystem for optical acquisition, tracking, and pointing to the receiver. The complexity of this hardware introduces concerns about cost, added weight, power consumption and reliability; the efforts needed to resolve these concerns have significantly delayed the implementation of laser communications systems on operational satellites. An example of the effects of this problem is the recent cancelation of the DSP Laser cross link program after 10 years of development and approximately \$600 million spent.

An alternate approach to laser communications which does not require extensive hardware on the satellite makes use of a retromodulator. The

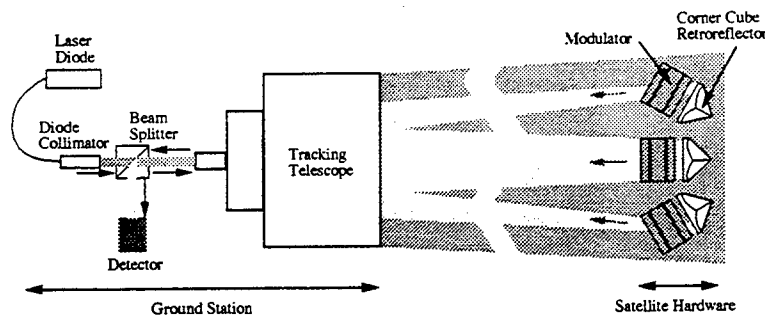


Figure 1: Conceptual diagram of a retromodulator communications system

alternate concept is as follows: First a laser site on the ground illuminates the satellite with either a continuous beam or a uniform train of pulses. On board the satellite is a corner cube, which reflects the incident laser light back to the transmitter on the ground. Associated with the corner cube is a modulation mechanism which reacts rapidly to change the intensity or the polarization of individual reflected pulses. By appropriately controlling this retromodulator the laser light pulses reflected back to the ground station are encoded with the communications signal (see Fig. 1).

Retromodulated communications offers the potential advantage of significantly reducing the size, power, and weight of hardware required aboard the satellite for laser communications while maintaining many of its advantages in general. This relatively new concept has many potential applications and may be appropriate where:

- The remote part of the communications link must be *very low power*.
- The remote site is inaccessible except by free space transmission.
- Weight and volume at the remote sight are a concern.
- There is a need for a "secure" data link.
- RF jamming is a concern.
- Simplicity at the remote site is desirable.
- A moderate data rate is required.

- Occasional delayed access to communications are acceptable.
- There is a need for simultaneous communications access to the remote sight.

A few specific applications that could benefit from such a technique are:

- Communications during atmospheric re-entry.
- Communications to/from downed pilots.
- Satellite to satellite links.
- Satellite to earth links:
 - Redundant or backup communications link.
 - Fly along noninterfering link.
 - Primary communications link.

Perhaps the most compelling reason for developing this type of communications system is to meet the severe power constraints of very small or micro satellites. This class of satellites typically do not have extensive deployable solar panels. The limited surface area of micro satellites for solar cells makes power a primary concern and techniques for conserving spacecraft power need to be explored.

Utah State University / Space Dynamics Lab has been involved in reflected laser communications research since 1992. Internal funding at USU allowed Dr. Swenson to initially study the concept with graduate students [1]. Dr. Charles Swenson and a graduate student spent 12 weeks during the summer of 1994 testing retromodulator concepts at Phillips Lab under the Air Force Office of Scientific Research Summer Research Program [2]. A flight opportunity was identified for a proof of concept test at the end of the summer. The Space Experiments directorate at Phillips Lab would be flying a high altitude balloon with an experiment for the Department of Energy in the spring of 1995. The primary mission is atmospheric sampling

but there was room for an additional payload providing it was small and can be easily mounted on the existing package. Development of the test hardware began in the fall of 1994 and successful ground testing of some of the retromodulators occurred at the Starfire Optical Range in December of 1994. Shortly after this testing the balloon flight for the DOE was cancelled ending the flight opportunity for the retromodulator experiment.

The Air Force Academy, through Mr. Gill Moore, became interested in this potentially low power communications technology and offered a test flight for hardware on one of the small "Falcon" satellites being built by Academy students. The Falcon satellites are to be placed into low Earth orbit by a launch system being developed from surplus Minuteman missile assets by the Space and Missile Test and Evaluation Directorate at Kirtland Air Force Base. As a step to the final satellite and good practice for ground based operations it was decided to conduct a test of the retromodulator laser communications package and as much of the Falcon satellite spacecraft as possible on a high altitude balloon over the Starfire Optical Range (SOR). Much of the hardware for the previously planned balloon test of 1995 could be used. Preparation for this new balloon test and the test itself were carried out during the 12 weeks of this year's Air Force Office of Scientific Research Summer Research Program.

This report presents an overview of the balloon launched retromodulator laser communications experiment activities conducted under this summer research program. This summer's successful balloon launch was a joint effort involving researchers from Utah State University working in PL/VTRA, the Starfire Optical Range (PL/LIG), the Air Force Academy, the Space and Missile Test and Evaluation Directorate, PL/SX, and Lawrence Livermore National Labs.

2 Balloon Retromodulator Experiment

The balloon reteromodulator experiment consisted of three principal parts. The first part is the actual the retromodulator experiment consisting of an array of retromodulators, drive electronics, an optical receiver, and a micro controller for operating the experiment. This portion of the experiment was built and tested by USU/SDL working out of PL/VTRA during the summer. The second part of the experiment is the ground station consisting of a laser transmitter, an optical receiver, and the pointing and tracking system. This was provided by SOR with some componets from USU/SDL and was also setup during the summer. The final component was the balloon system for carrying the retromodulator experiment to 105,000 ft altitude over the ground station. This portion was provided by the Air Force Academy with help from Norm Kjome of the University of Wyoming and Kjome Research. It consisted of a high altitude research balloon (approximately 70 feet in diameter and 50 feet tall at the float altitude) and support hardware for safty, tracking, and cut down.

A balloon at 24 km altitude is probably the next closest platform to an actual satellite for free space optical communications as most of the turbulent effects of the atmosphere would be observed. This program provided an ideal opportunity for a quick and relatively inexpensive demonstration of the reflected laser communications concept as well as a good "dry run" for a potential satellite demonstration. A brief outline of some of the objectives of the balloon test were:

1. Develop a working relationship between USU/SLD and the Air Force on reflected laser communications.
2. Demonstrate the basic principles and hardware of retromodulated laser

communications in a way that shows scaleability to a satellite.

3. Validate the communications link budget for retromodulated communications including atmospheric effects (scintillation, backscatter, etc) with vertical propagation through the atmosphere.
4. Test an assembly of FLC based retromodulators for parallel operation, temperature effects, speed, and modulation depth.

2.1 Ground Station

The 1.5 meter telescope at Starfire Optical range was used as the ground station consisting of a laser transmitter and a optical receiver. The transmitter consisted of A 5 watt 810 nm diode laser that was fiber coupled to a optical system mounted on the belly of the 1.5 meter telescope. These optics colimated the output of the fiber into about a 14 centimeter diameter beam with a half angle divergence of 1.5 mRad. The optical receiver consisted of the 1.5 meter telescope and a USU/SDL photomultiplier tube detector and pre-amp. SOR integrated all of these elements and developed new tracking capabilities for the 1.5 meter mount for high altitude balloons. Data from the photomultiplier tube was low-pass filtered at 10 kHz and digitized at 20 kHz for the experiment. Snatches of data were stored for post flight analysis.

2.2 Link Budget

Radio frequency communication links are typically described by engineers in terms of power, power gains, and power losses in decibel notation. Most engineers and scientists working with optical systems describe their systems in terms of number of photons, divergences, absorptions, and effective areas. The field of laser communications is mixed and we have chosen the latter system.

A detailed link budget was developed in spread sheet form for the balloon launch retromodulator experiment. This budget accounted for both signal and noise sources in the communications link and was based on a variety of measured and expected values. Signal to noise ratios were calculated for two types of detectors, a simple threshold system and an integrating system.

2.2.1 Radar Equation

The number of photons received, N_{rx} , by a ground station from a retroreflector at distance R with effective area A_{ret} obeys the well known radar equation [3]. A modified form of this equation which is descriptive of a retromodulator communication system is given by (1).

$$N_{rx} = \frac{N_{tx} A_{ret} A_{rx}}{\Omega_{tx} \Omega_{ret}} \tau_a \tau_o \quad (1)$$

Where N_{tx} is the number of photons transmitted into solid angle Ω_{tx} and received with effective area A_{rx} . The quantity Ω_{ret} describes the solid angle into which photons are returned from the retroreflector and accounts for intentional spoiling of the retroreflector and/or diffraction spreading. The global transmittance of the atmosphere and of the optics are accounted for by τ_a and τ_o respectively.

2.2.2 Link Calculations

The link budget calculations were made using a spread sheet on a PC. These calculations were made pre-flight and validated in sections using bench measurements, background observations, and ground based measurements of signal returns over a 5.4 km range. We present synopsis calculations for look angles of 45° from the ground station which corresponds to the longest

range and smallest crosssections for the retromodulators and for 0° (directly overhead), the best case condition.

Transmitter

<u>Design Element</u>	<u>Units</u>	<u>Value</u>
Link Wavelength	nm	810
Transmitter Power	Watts	5.000
Transmitter Photons Generated	photons/s	2.04E+19
Transmitter Divergence Half-Angle	Rad	1.5E-3
Solid Angle of Transmitter	Sr	7.07E-6
Transmitter Optics Losses	transmittance	0.657
Transmitter Photons Emitted	photons/s	1.34E+19

Receiver

<u>Design Element</u>	<u>Units</u>	<u>Value</u>
Antenna Diameter	m	1.5
Global Efficiency		0.8
Effective Area	m ²	1.75

Retro-reflector

<u>Design Element</u>	<u>Units</u>	<u>45°</u>	<u>0°</u>
Reflector Diameter	cm	2.54	
Look Angle to Array	Deg	45	0
Array Factor	unitless	0.167	2.082
Effective Area	cm ²	0.85	10.55
Optical Transmittance	transmittance	0.40	0.40
Spoiling half-angle	Rad	4.79E-4	
Solid Angle	Sr	7.21E-7	

Radar Equation

<u>Design Element</u>	<u>Units</u>	<u>45°</u>	<u>0°</u>
Maximum Altitude	m	32000	
Look Angle to Payload	Deg	45	0
Range to Retromodulator	m	45255	32000
Atmospheric Loss Outgoing	transmittance	0.7	
Atmospheric Loss Returning	transmittance	0.7	
Photons Received	photons/s	1.46E+7	7.25E+8

Noise

<u>Design Element</u>	<u>Units</u>	<u>Value</u>
Dark Noise Photons	photons/s	4.86E+4
Background Photons	photons/s	1.34E+5

Backscatter Photons	photons/s	1.53E+8	
Total Non-signal Photons	photons/s	1.53E+8	
<u>Signal</u>			
<u>Design Element</u>	<u>Units</u>	<u>Value</u>	
Modulation ratio	unitless	0.025	
Sample Period	s	5E-5	
Bit Period	s	5E-4	
Noise Factor	unitless	5	
<u>Detector</u>			
<u>Design Element</u>	<u>Units</u>	<u>45°</u>	<u>0°</u>
High Level Rate	photons/s	1.68E+8	8.79E+8
Low Level Rate	photons/s	1.54E+8	1.71E+8
Expected Std	photons/s	9.16E+6	2.10E+7
Signal/Noise (Threshold)	unitless	1.55	33.75
Samples/Bit		10	
Photons/Bit (High)	photons	8.39E+4	4.39E+5
Photons/Bit (Low)	photons	7.68E+4	8.57E+4
Signal/Noise (Integrating)	unitless	4.91	106.71
Error Rate	s ⁻¹	7.04E-3	~ 0

2.3 Balloon System

A 141,000 cubic ft polyethylene-film research-balloon was used to loft the retromodulator experiment and associated support and recovery equipment as illustrated in Fig. 2.3. The payload string included three strobe lights for detection by aircraft and SOR, radar transponders for the FAA, and GPS receivers and data links for position information. The balloon system was provided by the Air Force Academy with technical and equipment support from Norm Kjome of Kjome Research. The balloon and instrument string were launched from a site 2 miles west and 1 mile south of the Starfire Optical Range at the south end of Kirtland Air Force Base. Careful analysis of wind profile data, prior to launch, predicted that the balloon would travel east and north about 24 miles from SOR and then above 60,000 ft would move

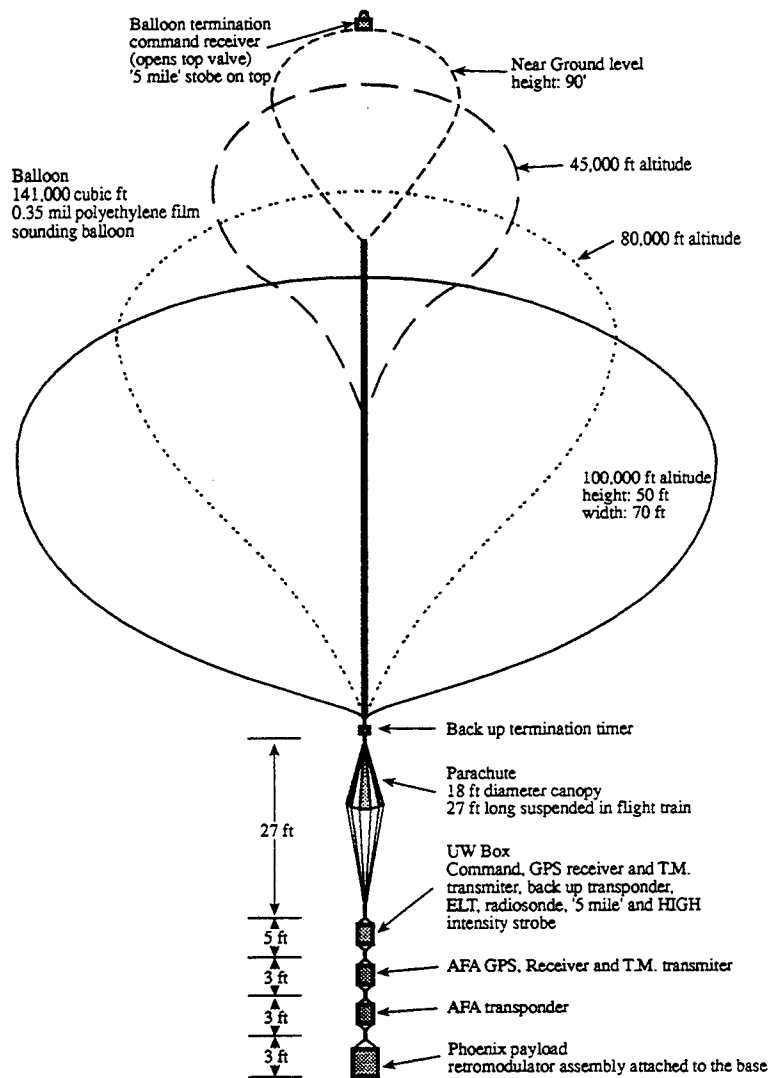


Figure 2: High altitude balloon and payload train.

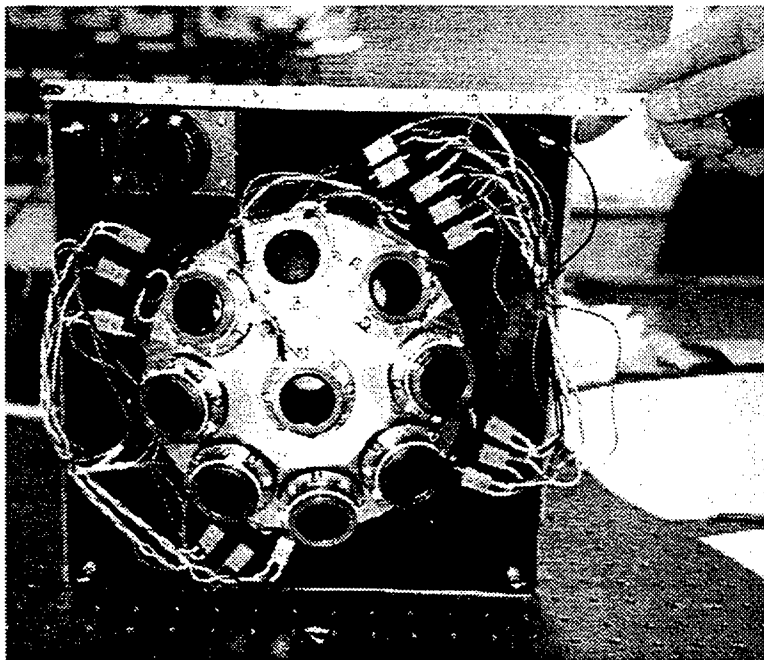


Figure 3: Preflight photograph of the retromodulator assembly

west, back over the ground station. The balloon followed this trajectory passing within about 10 degrees of vertical over SOR and continued about 30 miles west before it was brought down. This was the first high altitude balloon flight at SOR and was an important demonstrated of their tracking capability of this platform for near-space optical experiments.

3 Retromodulator Instrumentation

The retromodulator instrumentation consisted of an array of retromodulators as pictured in Fig. 3 along with control electronics. The electronics within the payload executed and monitored several tasks during the balloon flight. The on board micro controller provided modulating signals for the FLCs while measuring and controlling the temperature of each of the nine

individual retro assemblies. The hardware also contained the capability to receive optical uplink data sent from the ground station. A Tattletale model 8(TT8), from Onset Computer Corporation, managed these tasks. All of the electronic subsystems needed to operate the experiment interfaced with this advanced microcontroller. The flexibility, compact size and rich features of the TT8 made it a good choice for a on-board flight computer. The TT8 provided a Motorola 68332 microprocessor integrated with the following features: eight 12-bit analog to digital inputs, RS232 serial communication lines, high speed Queued Serial Module, multiple Time Processing Units, digital I/O lines. This computer consumes little power and operates off of a single power supply(7-15 V). A printed circuit board (PCB) was fabricated to interface the TT8 to the other modules. This board also implemented additional tools needed to run the experiment such as a 4 channel 12-bit and a 8 channel 8-bit digital to analog converters to drive the FLCs, high current switching transistors to control the heaters, an analog MUX to switch between temperature probes, and amplification circuitry for the optical uplink. A simple block diagram of the system can be seen in Fig. 3.

3.1 Retromodulator

The electrical components of the retromodulator assembly include the FLC, temperature probe, and heater. The balloon payload contained nine of each. The FLCs require a $\pm 5V$ square wave signal to drive them. This was accomplished by using one of the 12-bit D/A channels to produce the drive signal. Mixing this signal with nine other offset voltages, from the D/A chips, produced separate controls for each of the FLCs. This compensated for the offset biases of the op-amps and individual personalities of the liquid

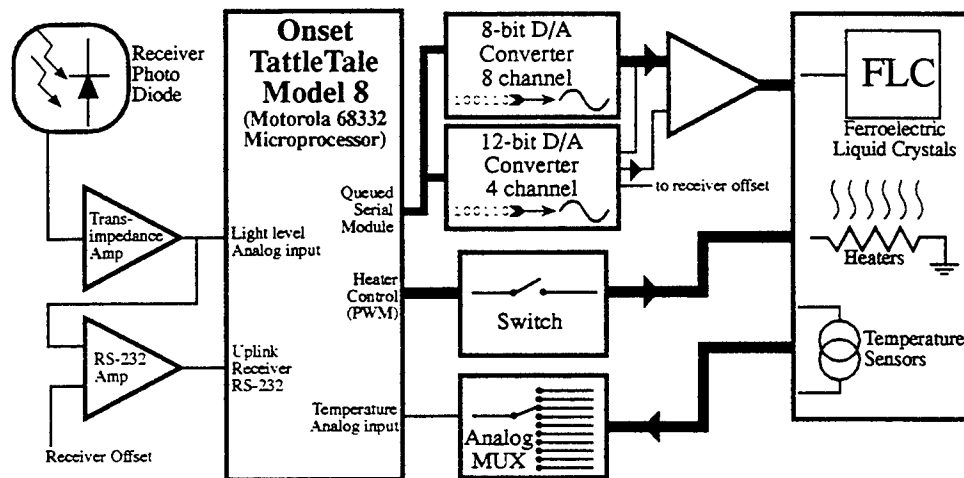


Figure 4: Block diagram of the electronics for the retromodulator communications experiment

crystals. Analog Devices AD590 temperature probes monitored the assemblies. The AD590 produces a current proportional to the temperature of the device. One microamp per degree Celsius ($1\mu\text{A}/^\circ\text{C}$). Dropping this current across a $10\text{k}\Omega$ resistor yielded ten millivolts per degree Celsius ($10\text{mV}/^\circ\text{C}$). To accommodate the nine temperature sensors an analog multiplexer was used. The output voltage of the MUX was input into one of the A/D channels on the TT8.

3.2 Optical Receiver

An optical detector system provided the balloon payload with the capability of receiving optical uplink data. A large active area silicon PIN photodiode was mounted on the bottom of the experiment. The two terminals of the unbiased diode fed up to the printed circuit board via a coaxial cable. Because of the low signal levels, careful considerations were made to reduce the additions of noise into the receiver system. The small currents from

the photo diode were input into a high gain transimpedance amplifier. The output of this stage connects to an A/D input of the TT8 and to another stage of gain. The A/D input allows the software to determine the level of light illuminating the experiment. This level is used to set the offset level in the next stage of gain. This light level is also valuable data because it can be used to verify part of the link analysis. The second stage of the receiver gain generates an RS-232 compatible signal. A simple level detection system converts the analog light level signal into a digital communications signal. An offset from one of the 12-bit D/A converter channels is used to determine the transition level. The gain of this stage is set high enough to ensure a strong RS-232 output even at low light levels. Basically this op-amp is operating as a comparator, driving the output to the rails of the device. The output connects to one of the RS-232 receiver ports on the TT8.

3.3 Programming/Data Logger port

The payload electronics implement an interface to other computers using either a RS-232 or RS-422 communication standards. This allows easy access to the microprocessor for the purpose of programming and debugging the hardware and software. During the balloon flight this port connected to a data logger that recorded system health information.

3.4 Software

During the balloon flight the on board computer ran software that controlled the experiment. This software was developed in C on a desktop PC and then cross compiled using the Aztec C compiler and then loaded into the flash memory of the Tattletale model 8. The main functions of the software were

to modulate the FLCs, control the temperatures of the retromodulator assemblies, receive uplink data optically from the ground station, and maintain a system timing so that the system would be synchronized with computers and clocks at the ground station. The FLCs were modulated through several test patterns in order to better prove the concept of a passive optical communications system. The majority of the data needed could be obtained by modulating the FLCs at a 500Hz square wave. Short bursts of higher frequency modulations also were tested. Because of the lack of an RF uplink or down link while the balloon was in flight, the software performed these tasks optically. Several times during each ten minute cycle, the computer would send down system health information by modulating the FLCs using a RS-232 like communications scheme. Using this method, it was possible to monitor system health information real time.

System Timing Schedule The system repeats every ten minutes Starting on the hour, ten past, twenty past, etc. The schedule is as follows:

2 Minutes 500 Hz Square wave modulating the FLCs on the balloon

30 Seconds System health down link (temperature, light level, etc.) This information is sent five times along with other test patterns during this thirty second period.

30 Seconds Higher Frequency tests. Six five-second bursts of high frequency square waves are modulating the FLCs on the balloon. The frequencies sent are 1k, 5k, 10k, 1k, 5k and 10k Hertz.

2 Minutes Stage 2, System health down link, is repeated four times. (i.e. the system health is sent 20 times during this stage.)

5 Minutes FLCs open. This allows for scintillation studies on the reflected laser beam.

Also written into the software but not used in flight was the ability to uplink data to the experiment. Information is sent to the balloon by modulating the outgoing laser. This data is received by the balloon payload and echoed by the FLCs on the balloon. Further development of this code will lead to the ability of real time uplink commands to be sent using only an optical link.

4 Flight and Preliminary Results

On Sunday night September 15, 1996 at 11:14 PM the balloon retromodulator experiment was successfully launched. One aborted attempt, due to high ground winds, had been made earlier in the week. The payload was illuminated by the 810 nm IR laser and optically tracked by telescopes at SOR though out the flight. Returns from the retromodulator were observed when the payload entered a 45° cone above the SOR. Data on-board the payload were transferred to the ground using only the passive retromodulators. Some preliminary of the data has been made and is shown in Fig. 4. A clear 500 Hz square wave can be observed with scintillation varying the amplitude in addition to some high frequency noise due to shot processes.

All technical objectives of the experiment were met and the researchers involved were very satisfied. The payload was available for data collection and experiments at the float altitude for over 1.5 hours. During which time snap shots of the received waveform were recorded for latter analysis. Approximately 140 M bytes of data is yet to be reduced and compared with the theoretical performance of this compunctions link.

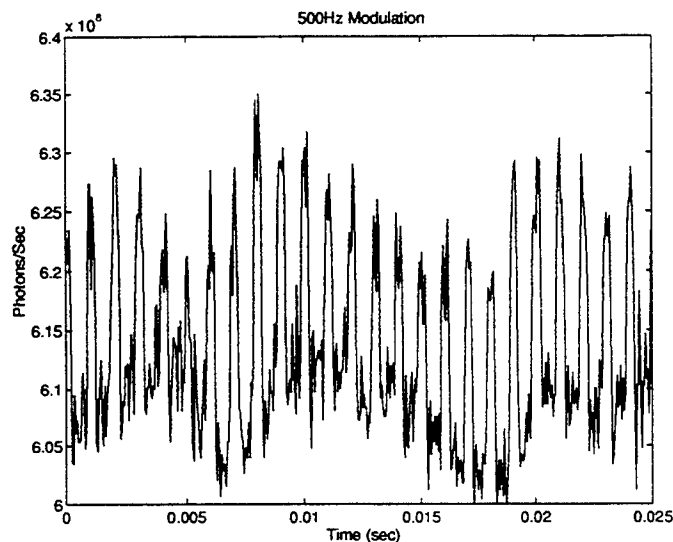


Figure 5: Sample data from a retromodulator at about 24 km range

5 Summary

The twelve weeks of the AFOSR summer research program have been very productive in developing and actually flying hardware for a retromodulator communications link. The tests achieved several firsts; The first retromodulator communications system to near space, the first laser communications experiment on a high altitude balloon, and the first high altitude balloon tracked by Starfire Optical Range. This project kept to an aggressively short time line and minimal budget. It involved a large number of people and several organizations including Utah State University/Space Dynamics Lab, the Air Force Phillips Lab (LI, VT, and SX), the Air Force Academy, and The Air Force Space and Missile Systems Center (TEB). All technical objectives of the experiment were met and the researchers involved were very satisfied. Considerable data has been generated that will help in the development of retromodulator communications systems. The AFOSR summer

research program is a very important and effective program in promoting cooperative research between Universities and the Air Force labs.

References

- [1] Gary Jensen and Charles Swenson. A laser downlink for small satellites using an optically modulated retroreflector. In *Proc. 6th Annual AIAA/USU Conf Small Sat.*, 1992.
- [2] Charles M. Swenson and Stephen Clarke. Reflected laser communication systems. Technical report, Air Force Office of Scientific Research / Phillips Lab, 1994.
- [3] Charles M. Swenson and Gary L. Jensen. Low-power optical transceiver for low earth orbit. In *SPIE Proceedings Vol 2553*, page 58, July 1995.

REGULARIZATION METHODS FOR LINEAR AND NONLINEAR RETRIEVAL PROBLEMS

Miguel Vélez-Reyes
Associate Professor
Laboratory for Applied Remote Sensing and Image Processing
Electrical and Computer Engineering Department

University of Puerto Rico Mayagüez Campus
P.O. Box 5000
Mayagüez, PR 00681-5000

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and
Phillips Laboratory

August 1996

REGULARIZATION METHODS
FOR LINEAR AND NONLINEAR RETRIEVAL PROBLEMS

Miguel Vélez-Reyes
Associate Professor
Electrical and Computer Engineering Department
University of Puerto Rico Mayagüez Campus

Abstract

Retrieval of atmospheric and surface characteristics, such as atmospheric temperature and water vapor content, from remotely sensed data is at the heart of applications in meteorology, atmospheric, oceanography, and geophysical sciences. The relation between features of interest and the measured radiance is given by the radiative transfer equation. The retrieval problem is then equivalent to the inversion of the radiative transfer equation. Many of these inversion problems are ill-posed and therefore regularization techniques are needed to convert the retrieval problem into a well-posed problem that can be numerically solved. In this report, some results regarding the application of regularization techniques to the temperature retrieval problem from microwave radiometry near O₂ absorption band at 60 GHz are presented. Results show that regularization methods can be capable of determining the tropopause height independent of the initial guess with need of minimal prior statistical information about sensor noise and profile distribution.

REGULARIZATION OF LINEAR AND NONLINEAR RETRIEVAL PROBLEMS

Miguel Vélez-Reyes

1 Introduction

Passive remote sensing of the atmosphere is used to determine the atmospheric state. A radiometer measures microwave emissions from earth's atmosphere and surface. The radiance measured by the radiometer is proportional to the brightness temperature. This brightness temperature can be used to estimate atmospheric parameters such as temperature and water vapor content. These quantities are of primary importance for different applications in meteorology, oceanography, and geophysical sciences.

Depending on the range in the electromagnetic spectrum being measured by the radiometer and the atmospheric quantities to be estimated, the retrieval or inverse problem of determining atmospheric parameters from brightness temperature might be linear or nonlinear. In most applications, the retrieval problem requires the inversion of a Fredholm integral equation of the first kind making this an ill-posed problem. The numerical solution of the retrieval problem requires the transformation of the continuous problem into a discrete problem. The ill-posedness of the continuous problem translates into ill-conditioning of the discrete problem. Another difficulty associated with the retrieval problem is that they are underconstrained because there are only few brightness temperatures being measured by the sensor while the quantities of interest are estimated at a higher vertical resolution (more unknowns). To further constraint the problem, prior information is incorporated into the problem to convert the ill-posed problem into a well-posed one.

In this report, we present the results of our work in applying different regularization techniques to the discrete problem associated with temperature retrievals using brightness temperatures from the SSM/T-1 sensor. Simulation results are presented which show the potential of these techniques to improve the performance of retrieval. In particular, no statistical assumptions are needed and the algorithms were capable of correctly estimating the temperature profile at the tropopause independent of the initial guess.

2 Radiative Transfer Theory in the Microwave Region

Radiative transfer theory describes the intensity of radiation propagating in a general class of media that absorbs, emit, and scatter the radiation [14]. The radiative transfer equation for a plane-parallel atmosphere is given by

$$\cos \theta \frac{d I_{\nu}}{dz} = -\sigma(z) I_{\nu} + J_{\nu}(z) \quad (1)$$

where $I_\nu(z)$ is the instantaneous radiant intensity that flows at each point in the medium per unit area, per unit of solid angle, at a given frequency ν ; $\sigma(z)$ is the extinction coefficient; and J is a source term. These last two quantities describe the loss/gain into the given direction. The angle θ is the direction angle with respect to the vertical axis z with $\theta = 0$ when pointing upwards.

In the general case, scattering into and from other directions can lead to both gains and losses to the intensity and are taken care by the terms σ and J . For the microwave region, the scattering term is usually neglected [10]. If scattering is neglected, the only source term to consider is that due to local emission and the extinction coefficient reduces to the absorption coefficient σ_a . Assuming local thermodynamic equilibrium, each point can be characterized by a temperature T and from Kirchoff's law we get

$$J_\nu(z) = \sigma_a(z) B_\nu(T(z)) \quad (2)$$

where $B_\nu(T)$ is the Planck function:

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp \frac{h\nu}{kT} - 1} \quad (3)$$

where $h = 6.625 \times 10^{-34}$ Js is Planck's constant, $c = 2.988 \times 10^8$ m/s is the speed of light, and $k = 1.381 \times 10^{-23}$ J/K.

Equation (1) is a linear non-homogeneous first-order differential equation with solution

$$I_\nu(z) = I_\nu(z_o) e^{-\sec \theta \int_{z_o}^z \sigma_a(z) dz} + \sec \theta \int_{z_o}^z \sigma_a(z) e^{-\sec \theta \int_\gamma^z \sigma_a(\xi) d\xi} B_\nu(T(\gamma)) d\gamma \quad (4)$$

where $I_\nu(z_o)$ is the boundary condition. The boundary condition will be treated later.

In the microwave region of the spectrum

$$h\nu \ll kT$$

which results in (3) taking the form

$$B_\nu(T) = \frac{2h\nu^2 kT}{c^2} = \frac{2kT}{\lambda^2} \quad (5)$$

where λ is the wavelength (m). This is known as the Raleigh-Jeans approximation. From this expression, it is clear that in the microwave region the energy emitted is proportional to the physical temperature T . Another commonly used result from this relation is to define a scaling of the intensity I_ν as follows

$$T_b = \frac{\lambda^2}{2k} I_\nu \quad (6)$$

The quantity $T_b(\nu)$ is called the *brightness temperature* which is commonly used in the microwave retrieval literature instead of I_ν . In terms of brightness temperature and using (5), (4) takes the form

$$T_b(z) = T_b(z_o) e^{-\sec \theta \int_{z_o}^z \sigma_a(z) dz} + \sec \theta \int_{z_o}^z \sigma_a(z) e^{-\sec \theta \int_\gamma^z \sigma_a(\xi) d\xi} T(\gamma) d\gamma \quad (7)$$

Here the dependency of all these quantities in frequency ν is not shown for convenience.

A convenient way to represent (7) is in terms of the optical thickness defined by

$$\delta(z) = \int_z^\infty \sigma_a(\xi) d\xi \quad (8)$$

here ∞ represents the top of the atmosphere (TOA). Notice also that

$$d\delta = -\sigma_a(z) dz$$

In terms of optical thickness, (7) can be rewritten as

$$T_b(z) = T_b(z_o)e^{-(\delta(z_o)-\delta(z))\sec\theta} + \sec\theta \int_{z_o}^z \sigma_a(\gamma)e^{-(\delta(\gamma)-\delta(z))\sec\theta} T(\gamma) d\gamma \quad (9)$$

For our purpose, it is of interest to solve this equation to obtain the brightness temperature that a satellite will measure at the top of the atmosphere when looking to the surface at an angle θ off nadir. This will correspond to $z_o = 0$ surface (sfc) and $z = \infty$ in (9). For this case, the quantities contributing to the boundary condition at the surface are illustrated in Figure 1. The boundary term $T_b(0)$ is given by

$$T_b(0) = \epsilon T_s + (1 - \epsilon) T_d$$

where T_s is the surface temperature, T_d is the downwelling radiation reflected by the surface back towards the satellite, and ϵ is the surface emissivity. For the reflected component, it is assumed that the surface is a smooth, homogeneous, and isothermal so only the radiation in the specular direction θ is accounted for. In our reference coordinates, the propagation angle for the downwelling radiation is $\pi - \theta$.

The downwelling radiation is obtained by solving the radiative transfer equation where integration is from the TOA to the surface with propagation angle $\pi - \theta$. This is taken care in (7) by setting $z_o = \infty$ and $z = 0$. The boundary term in this case is given by the cosmic microwave background emission with $T_c = 2.7$ K. The downwelling radiation is then given by

$$T_b(z) = T_b(z_o)e^{-\sec\theta \int_{z_o}^0 \sigma_a(z) dz} + \sec\theta \int_{z_o}^0 \sigma_a(z)e^{-\sec\theta \int_z^0 \sigma_a(\xi) d\xi} T(z) dz \quad (10)$$

To substitute the downwelling radiation contribution in (9), it is more convenient to rewrite (10) as follows

$$T_d = T_c e^{-\delta(0)\sec\theta} + \int_0^\infty T(z) e^{-(\delta(0)-\delta(z))\sec\theta} \sigma_a(z) dz \quad (11)$$

Substituting this expression into (9) and combining the terms in the integrals, we obtain the expression for the brightness temperature at the TOA at an angle θ off nadir

$$T_b = \left[\epsilon T_s + (1 - \epsilon) T_c e^{-\delta(0)\sec\theta} \right] e^{-\delta(0)\sec\theta} + \int_0^\infty T(z) \left[1 + (1 - \epsilon) e^{-2(\delta(0)-\delta(z))\sec\theta} \right] e^{-\delta(z)\sec\theta} dz \quad (12)$$

Simulation of this expression if all quantities were known is a simple matter. However, the computation of the optical thickness and the absorption parameter requires the use of databases containing information about the spectral characteristics of atmospheric constituents such as HITRAN. In our work, all optical depth computations were carried out using the FASE Radiative Transfer Code.

3 Atmospheric Remote Sensing

The satellite instrument measures radiance that arrives into its field of view. The radiance that arrives is the sum of the radiance emitted and reflected by the surface, emitted and reflected by the atmosphere, and that scattered by the atmosphere into the field of view of the instrument. The relative contribution of each component depends on the region of the spectra seen by the instrument. Remote sensing of the surface takes advantage of those regions of the spectrum where the atmosphere is transparent or nearly so. In the case of atmospheric remote sensing, the satellite sensor is looking at regions in the spectra where the atmosphere blocks the radiance emitted or reflected by the surface and therefore it receives that radiation that is emitted or reflected by the atmosphere.

The interaction of electromagnetic waves with the atmosphere depend on the characteristics of the propagating wave (primarily its wavelength), the physical characteristics of the atmosphere and its constituents (pressure, temperature, density, absorbing gases, suspended particles). The mechanisms for interactions are: scattering, absorption, emission, and refraction. In regions where the atmospheric constituents characteristics are known or understood as in the 60 GHz oxygen and 183 GHz water vapor absorption lines. Measurements of brightness temperature can be used to infer atmospheric properties of interest.

The relation between atmospheric properties with brightness temperature is given by the radiative transfer equation discussed earlier. Therefore the problem of interest is to infer the atmospheric quantities of interest from measured brightness temperature by inversion of the radiative transfer equation.

4 Temperature Retrieval Problem

If the atmosphere strongly absorbs, most of the contribution to the measured brightness temperature will come from the atmosphere itself. In the case of the microwave region of the spectrum, scattering is negligible and the energy into the field of view of the sensor will come from atmospheric emission. Assuming that the satellite is looking at nadir (i.e. $\theta = 0$), and that the surface temperature T_s and emissivity ϵ are known, (12) can be rewritten as

$$T_b - [\epsilon T_s + (1 - \epsilon) T_c e^{-\delta(0)}] e^{-\delta(0)} = \int_0^\infty T(z) [1 + (1 - \epsilon) e^{-2(\delta(0) - \delta(z))}] e^{-\delta(z)} dz \quad (13)$$

or equivalently as

$$\tilde{T}_b = \int_0^\infty T(z) \frac{dS(z)}{dz} dz \quad (14)$$

where

$$\begin{aligned} S(z) &= [1 - (1 - \epsilon) e^{-2(\delta(0) - \delta(z))}] e^{-\delta(z)} \\ \tilde{T}_b &= T_b - [\epsilon T_s + (1 - \epsilon) T_c e^{-\delta(0)}] e^{-\delta(0)} \end{aligned}$$

Channel	Peaking Height (Km)	Frequency (GHz)	Bandwidth (MHz)
1	0	50.5	400
2	2	53.2	400
3	6	54.35	400
4	10	54.9	400
5	30	58.4	115
6	16	58.825	400
7	22	59.4	250

Table 1: SSM/T-1 Channel Description.

If the absorber is uniformly mixed with a known concentration, as O_2 , the quantity

$$K(\nu, z) = \frac{dS(\nu, z)}{dz}$$

is known and the temperature profile $T(z)$ could be retrieved by inverting (14). The function $K(\nu, z)$ is called in the literature [10] the *weighting function*. From this point on, in our discussion we would not distinguish between T_b and \tilde{T}_b in (14).

O_2 has several absorption lines around between 50 and 60 GHz. The Special Sensor Microwave Temperature-1 (SSM/T-1) sensor of the sensor suite of the DMSP satellite has 7 channels located in the 50 to 60 GHz range used for temperature retrievals. A summary of the SSM/T-1 sensor characteristics is given in Table 1 from [15]. The weighting functions for the SSM/T-1 channels for a U.S. Standard Atmosphere are shown in Figure 2. A particular characteristic of these functions are their peaks at different heights basically this establishes that the brightness temperature is more sensitive to the temperature at a particular height. In the O_2 band, the shape of the weighting function is independent of the temperature making the inversion of (14) a linear inversion problem.

4.1 Ill-Posedness of Temperature Retrievals

Equation (14) is a Fredholm integral equation of the first kind.

$$g(x) = \int_a^b K(x, s)u(s)ds \quad (15)$$

The inversion of a Fredholm equation is an ill-posed problem [7, 20]. To understand what an ill-posed problem is, we need first to define what is a well-posed problem. The following definition is due to Hadamard [7, 20].

Definition 4.1 Well-Posed Problem

A problem is given by a datum g and a solution u , and it is well-posed (in the sense of Hadamard) when

1. *For each datum g in a class of functions \mathcal{Y} there exists a solution u in a prescribed class \mathcal{X} (existence).*
2. *The solution u is unique in \mathcal{X} (uniqueness).*

3. The dependence of u upon g is continuous.

An *ill-posed* problem is a problem that is not *well-posed* (i.e. it fails at least one of the conditions stated above).

The ill-posedness of the inversion of the Fredholm integral equation is shown by the Riemman-Lebesgue lemma [7] which states that if $K(\cdot, \cdot)$ is a square integrable kernel, then

$$\int_0^\pi K(x, s) \sin ns \, ds \rightarrow 0 \text{ as } n \rightarrow \infty$$

Therefore, for large values of n , the slightly perturbed data

$$\tilde{g}(x) = g(x) + A \int_0^\pi K(x, s) \sin ns \, ds$$

corresponds to a solution $u(s) + A \sin ns$ that differs significantly from $u(s)$. This implies that solutions to the Fredholm equation depend discontinuously on the data. To solve this problem prior information about the nature of the solution is brought in to convert the ill-posed problem into a well-posed one. That is, the problem is constrained to create a problem with a unique solution.

4.2 Discrete Problem

To numerically solve the temperature retrieval problem, first we discretize (14) by approximating the integral with a numerical integration formula.

$$\tilde{T}_b(\nu) = \int_0^\infty T(z) \frac{dS(\nu, z)}{dz} dz \approx \sum_{i=0}^n w_i K(\nu, z_i) T_i \quad (16)$$

where w_i is a quadrature weight and T_i is equal to $T(z_i)$. This results in the algebraic linear system of equation

$$\mathbf{T}_b = \mathbf{K}\mathbf{T} + \mathbf{e} \quad (17)$$

where $\mathbf{T}_b \in \mathcal{R}^m$ and $\mathbf{T} \in \mathcal{R}^n$ are the brightness and atmospheric temperature vector; $\mathbf{K} \in \mathcal{R}^{m \times n}$ is the matrix of weighting functions; and \mathbf{e} is an error term associated with measurement noise and the truncation error arising from the discretization of the integral equation. The number of measured brightness temperatures m is usually smaller than the vertical resolution or number of temperature levels n to estimate. In the case of the SSM/T-1 sensor, there are $m = 7$ channels and normally $n = 20$ temperature levels. Therefore the resulting algebraic linear system of equations (17) is underconstrained (i.e. there are more unknowns than equations).

4.3 Regularization of the Discrete Problem

The temperature retrieval problem is related to the solution of the linear system of equations (17). This problem has two major difficulties associated with it:

- ill-conditioning due to the ill-posedness of the associated integral equation, and
- multiple solutions because of trying to estimate more temperature levels than measurements available.

4.3.1 Characterization of the Solution Set

In the following discussion, we will assume that $\text{rank}(\mathbf{K}) = m$, as is the case for the SSM/T-1 sensor. To characterize the solution set for (17), we will need the following quantities

Definition 4.2 Moore-Penrose Pseudoinverse: Let \mathbf{K} be a $m \times n$ matrix with $m < n$ such that $\text{rank}(\mathbf{K}) = m$, then its pseudo-inverse is given by

$$\mathbf{K}^\dagger = \mathbf{K}^T (\mathbf{K}\mathbf{K}^T)^{-1}$$

Notice that $\mathbf{K}\mathbf{K}^\dagger = \mathbf{I}_{m \times m}$.

Lemma 4.1 Let $\mathcal{R}(\mathbf{K})$ and $\mathcal{N}(\mathbf{K})$ denote the range and null spaces of the matrix \mathbf{K} respectively, then

1. $\mathbf{P}_K = \mathbf{K}\mathbf{K}^\dagger$ is the orthogonal projector onto $\mathcal{R}(\mathbf{K})$,
2. $\mathbf{Q}_K = \mathbf{I} - \mathbf{P}_K$ is the orthogonal projector onto the complement of $\mathcal{R}(\mathbf{K})$,
3. $\mathbf{P}_{K^T} = \mathbf{K}^\dagger \mathbf{K}$ is the orthogonal projector onto $\mathcal{R}(\mathbf{K}^T)$,
4. $\mathbf{I} - \mathbf{P}_{K^T}$ is the orthogonal projector onto $\mathcal{N}(\mathbf{K})$.

Proof: see [18]. □

The following theorem gives a complete characterization of the solution set of (17) [18].

Lemma 4.2 (Penrose) The solution of the problem (17) has the general form

$$\mathbf{T} = \mathbf{K}^\dagger \mathbf{T}_b + \mathbf{Q}_{K^T} \mathbf{z} \quad (18)$$

where $\mathbf{z} \in \mathcal{R}^m$ is arbitrary. Of all solutions, $\mathbf{K}^\dagger \mathbf{T}_b$ uniquely has the smallest 2-norm.

The second component of (18) corresponds to a component in the null space of \mathbf{K} . The dimension of the null-space of \mathbf{K} is $(n-m)$ so there are $(n-m)$ degrees of freedom in selecting the solution to (17). Therefore, there is no unique solution to the temperature retrieval problem.

To overcome this difficulty, we will use the so called regularization methods. Regularization theory [20] transforms an ill-posed problem to a well-posed one, using *a priori* knowledge on the nature of the solution. Depending on the prior information, regularization techniques can be classified into two major groups: statistical and deterministic.

4.3.2 Statistical Regularization

As we have seen, there is no unique solution to the temperature retrieval problem. In statistical regularization, prior statistical information is used to regularize the temperature retrieval problem. Our prior information in this case is the prior distribution of the temperature profiles $p_T(T)$ and the conditional distribution $p_{T_b/T}(T_b)$.

According to Bayesian estimation theory [11], the best estimator $\hat{\mathbf{T}}$ based on the brightness temperature observation \mathbf{T}_b of the temperature profile \mathbf{T} in a mean square sense is defined by minimization

$$\hat{\mathbf{T}} = \arg \min_{\mathbf{T} \in \mathcal{R}^n} \int_{-\infty}^{\infty} \|\mathbf{T} - E(\mathbf{T}/\mathbf{T}_b)\|^2 p_{T/T_b}(\mathbf{T}) d\mathbf{T} \quad (19)$$

and the best estimate is given by the conditional mean

$$\hat{\mathbf{T}} = E(\mathbf{T}/\mathbf{T}_b) \quad (20)$$

We will refer to this estimator as the minimum mean square estimator (MMSE). The analytical determination of this function might be a very difficult task. In many instances, the estimator is constrained to be linear which results in the Linear Minimum Mean Squares Estimator (LMMSE) [11]

$$\hat{\mathbf{T}} = \bar{\mathbf{T}} + \mathbf{A}_{T,T_b} \mathbf{A}_{T_b,T_b}^{-1} (\mathbf{T}_b - \bar{\mathbf{T}}_b) \quad (21)$$

where $\bar{\mathbf{T}}$ is the a priori mean of \mathbf{T} , $\bar{\mathbf{T}}_b$ is the mean of \mathbf{T}_b given \mathbf{T} , \mathbf{A}_{T,T_b} is the cross covariance between \mathbf{T} and \mathbf{T}_b , and \mathbf{A}_{T_b,T_b} is the conditional covariance of the brightness temperature. The LMMSE is easy to construct, since only the first and second order statistics are needed rather than their complete probability densities. Also, if \mathbf{T} and \mathbf{T}_b are jointly Gaussian, the LMMSE is the optimal Bayesian MMSE.

Since \mathbf{T} and \mathbf{T}_b are linearly related by (17), the computation of the first and second order statistics are given by

$$\begin{aligned} \bar{\mathbf{T}}_b &= \mathbf{K} \bar{\mathbf{T}} \\ \mathbf{A}_{T,T_b} &= (\mathbf{K}^T \mathbf{A}_e \mathbf{K} + \mathbf{A}_T)^{-1} \mathbf{K}^T \\ \mathbf{A}_{T_b,T_b} &= \mathbf{A}_e \end{aligned}$$

using these definitions, equation (21) can be rewritten as

$$(\mathbf{K}^T \mathbf{A}_e \mathbf{K} + \mathbf{A}_T) (\hat{\mathbf{T}} - \bar{\mathbf{T}}) = \mathbf{K}^T \mathbf{A}_e^{-1} (\mathbf{T}_b - \mathbf{K} \bar{\mathbf{T}}) \quad (22)$$

This representation will be of particular importance when comparing different regularization methods. The a priori statistical information needed is estimated from archived radiosonde data.

4.4 Deterministic Methods

In this section, we will look at two regularization methods for ill-posed linear algebraic systems of equations: Tikhonov regularization, and discrepancy principle regularization. Other methods are discussed in [8, 9].

4.4.1 Tikhonov's Regularization

One way to regularize (17) is computing $\hat{\mathbf{T}}$ as the solution to the optimization problem

$$\hat{\mathbf{T}}_\lambda = \arg \min_{\mathbf{T} \in \mathcal{R}^n} \left[\|\mathbf{KT} - \mathbf{T}_b\|^2 + \lambda^2 \|\mathbf{L}(\mathbf{T} - \mathbf{T}_o)\|^2 \right] \quad (23)$$

where \mathbf{T}_o is a prior temperature profile estimate, $\|\cdot\|$ is the 2-norm¹, and λ is the regularization parameter. Notice that this optimization problem can be rewritten as the linear least squares problem

$$\hat{\mathbf{T}}_\lambda = \arg \min_{\mathbf{T} \in \mathcal{R}^n} \left\| \begin{bmatrix} \mathbf{K} \\ \lambda \mathbf{L} \end{bmatrix} \mathbf{T} - \begin{bmatrix} \mathbf{T}_b \\ \lambda \mathbf{LT}_o \end{bmatrix} \right\|^2 \quad (24)$$

and it has a unique solution $\hat{\mathbf{T}}_\lambda$ if the null spaces of \mathbf{K} and \mathbf{L} intersect trivially at $\mathbf{T} = \mathbf{0}$. Efficient algorithms for the solution of this problem are described in [4].

A key issue in this method is the selection of the regularization parameter λ . The value used in the simulation results presented here was based on the L-curve method described in [8]. The optimal value of λ balances the prediction error $\|\mathbf{KT} - \mathbf{T}_b\|$ with the regularization error $\|\mathbf{L}(\mathbf{T} - \mathbf{T}_o)\|$.

A useful formulation of the solution to (23) is in terms of the normal equations for (24) given by

$$(\mathbf{K}^T \mathbf{K} + \lambda^2 \mathbf{L}^T \mathbf{L}) (\hat{\mathbf{T}}_\lambda - \mathbf{T}_o) = \mathbf{K}^T (\mathbf{T}_b - \mathbf{KT}_o) \quad (25)$$

This formulation has some similarities with (22). As we can see regularization in (22) is achieved by adding a positive definite matrix \mathbf{A}_T to the positive semidefinite matrix $\mathbf{K}^T \mathbf{K}$ while here it is achieved by adding the positive definite matrix $\lambda^2 \mathbf{L}^T \mathbf{L}$. We will look into this issues in more detail later.

4.4.2 Discrepancy Method

Another possibility to regulate the temperature retrieval problem is by computing $\hat{\mathbf{T}}$ as the solution to the quadratically constrained linear least squares problem

$$\begin{aligned} \hat{\mathbf{T}}_\alpha &= \arg \min_{\mathbf{T} \in \mathcal{R}^n} \|\mathbf{L}(\mathbf{T} - \mathbf{T}_o)\|^2 \\ \text{subject to } &\|\mathbf{KT} - \mathbf{T}_b\|^2 \leq \alpha^2 \end{aligned} \quad (26)$$

where α plays the role of a regularization parameter. The solution to this problem can be made identical to $\hat{\mathbf{T}}_\lambda$ for a suitably chosen α [8, 9]. We prefer to select α based on the measurement noise norm $\|\mathbf{e}\|$.

Using Lagrange multiplier theory [1], it can be shown that the estimated temperature profile satisfies

$$\begin{aligned} \left(\mathbf{K}^T \mathbf{K} + \frac{1}{\mu^2} \mathbf{L}^T \mathbf{L} \right) (\hat{\mathbf{T}}_\alpha - \mathbf{T}_o) &= \mathbf{K}^T (\mathbf{T}_b - \mathbf{KT}_o) \\ \|\mathbf{K}\hat{\mathbf{T}}_\alpha - \mathbf{T}_b\|^2 &= \alpha^2 \end{aligned} \quad (27)$$

for some lagrange multiplier μ . Here, we wanted to point out the appearance of another regularized normal equation (27) similar to (22) and (25).

¹The 2-norm is given by $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$.

4.5 Nomal Equations in Regularization

In all three methods discussed so far, the computation of the regularized solution involves the solution of a system of equations of the form

$$(A^T A + \lambda^2 B^T B) (\hat{T}_\lambda - T_o) = B^T (T_b - K T_o) \quad (28)$$

and the retrieved temperature profile is given by

$$\hat{T}_\lambda = T_o + (A^T A + \lambda^2 B^T B)^{-1} B^T (T_b - K T_o) \quad (29)$$

Notice that regularization is achieved by adding a positive definite matrix $\lambda B^T B$ to the rank deficient matrix $A^T A$.

A convenient way to see the effect of the regularization is using the generalized singular value decomposition of the matrix pair (A, B) [6, 16]. The following definition of the GSVD has been modified to fit our problem.

Definition 4.3 Generalized Singular Value Decomposition: Let $A \in \mathcal{R}^{m \times n}$, with $m \leq n$ and $B \in \mathcal{R}^{n \times n}$ with $\text{rank}(B)=n$, then there exist orthogonal $U \in \mathcal{R}^{m \times m}$ and $V \in \mathcal{R}^{n \times n}$ and an invertible $X \in \mathcal{R}^{n \times n}$ such that

$$U^T A X = [\Sigma, 0] \quad (30)$$

and

$$V^T B X = \begin{bmatrix} M & 0 \\ 0 & I_{n-m} \end{bmatrix} \quad (31)$$

where

$$\Sigma = \text{diag} \{ \sigma_1, \dots, \sigma_m \}$$

$$M = \text{diag} \{ \mu_1, \dots, \mu_m \}$$

and

$$1 \geq \sigma_1 \geq \dots \geq \sigma_m > 0 \quad 0 \leq \mu_1 \leq \dots \leq \mu_m \leq 1$$

They are normalized such that

$$\sigma_i^2 + \mu_i^2 = 1, \quad i = 1, 2, \dots, m$$

Then the generalized singular values γ_i of (A, B) are defined as the ratios

$$\gamma_i = \frac{\sigma_i}{\mu_i} \quad (32)$$

Using the GSVD, (29) can be rewritten as

$$\hat{T}_\lambda - T_o = \sum_{i=0}^m f_i x_i \left(\frac{v_i^T (T_b - K T_o)}{\sigma_i} \right) \quad (33)$$

where

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2 \mu_i^2} = \frac{\gamma_i^2}{\gamma_i^2 + \lambda^2} \quad (34)$$

The fraction f_i is called a filter factor.

In the case of the ordinary singular value decomposition (e.g. $\mathbf{B} = \mathbf{I}$), (33) takes the form

$$\hat{\mathbf{T}}_\lambda - \mathbf{T}_o = \sum_{i=0}^m f_i \mathbf{x}_i \left(\frac{\mathbf{v}_i^T (\mathbf{T}_b - \mathbf{K} \mathbf{T}_o)}{\sigma_i} \right) \quad (35)$$

with

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \quad (36)$$

Here, we can see how the regularization helps to stabilize the solution for the retrieval problem. If the regularization factor λ is selected to be larger than the smaller singular value of \mathbf{A} then $f_i < 1$ and if $f_i \rightarrow 0$ faster than σ_i , the contribution to the solution by $\mathbf{v}_i^T (\mathbf{T}_b - \mathbf{K} \mathbf{T}_o) / \sigma_i$ associated of a small singular values is effectively filtered out. The same idea can be extended to generalized singular values [8, 9].

5 Simulation Experiments

In this section, we present some simulation results that illustrate the use of deterministic methods to regularize the temperature retrieval problem. For the simulation experiments, a 98 layer atmosphere based on the US Standard Atmosphere model was used. The optical thickness for each layer was computed using the FASE code. The radiative transfer equation was numerically integrated using the trapezoidal rule. In our simulations, the surface emissivity ϵ was set to 0.9 and the surface temperature T_s set to 288.2 degrees Kelvin.

Figures 3 to 6 show the results of applying Tikhonov's regularization to the temperature retrieval problem with perfect measurements (no noise). The regularization parameter λ was set 0.0044 that is equal to the smallest singular value of \mathbf{K} putting little or no regulation on the problem. The solid line is the retrieved temperature profile, the dashed line is the actual temperature profile, and the dash-dot line is the initial guess fed to the algorithm. We can see from this simulations that the algorithm was capable of estimating temperature up to 40 km. Beyond 40 km, the resulting estimate was identical to the initial guess. This can be easily seen from the sensitivity functions shown in Figure 2 that are almost equal to zero beyond that height. An important result is how the retrieval algorithm is capable of determining the height of tropopause when fed with initial guesses that have the tropopause height far from the actual height as shown in Figures 3, 5 and 6. The importance of determining this peak comes from the fact that most important weather features are located at this region of the atmosphere. Also, the location of the tropopause peak serves as a figure of merit in evaluating the performance of temperature retrieval algorithms. We are not showing results for the algorithm based on the discrepancy principle regularization since they were identical to those of Tikhonov regularization.

Figures 7 to 10 show the performance of Tikhonov regularization under the presence of noisy data. The noise vector added to the measured brightness temperature has a normal distribution with zero mean and unit variance. The actual numbers added to each brightness temperature are given by

$$\mathbf{e} = [1.7971 \quad 0.2641 \quad 0.8717 \quad -1.4462 \quad -0.7012 \quad 1.2460 \quad -0.6390]^T$$

The regularization parameter for this case was $\lambda = 0.016$ that regulates the effect that the oscillatory right singular vector of \mathbf{K} associated with the small singular value at 0.0044 has on the solution. The filter factor for the small singular value is about 0.25. Notice that the noise causes the retrieved profile to be noisier with a maximum error in the first 20 km of 10 degrees Kelvin. Quite high compared to some retrieval methods that claim accuracies of 0.5 degrees Kelvin. However, the algorithm is still capable of determining the height of the tropopause.

Figures 11 to 14 show the results for the retrievals in the noisy case computed using discrepancy principle regularization. The regularization parameter for this case was set at $\alpha = 3$. That value corresponds to three standard deviations of the noise distribution. The resulting retrievals are smoother than those retrievals from Tikhonov regularization. The maximum error is in the neighborhood of 3 to 4 degrees kelvin in the tropopause. The estimation of the location of the tropopause peak is also improved.

6 Nonlinear Retrievals

Nonlinear retrieval problems occur when trying to retrieve atmospheric constituents from brightness temperature measurements. An important example of such problems is the retrieval of water vapor using data measured around the 183 GHz water vapor line [12, 13, 17, 22]. This is the range covered by the SSM/T-2 microwave radiometer on board of the DMSP satellite. Also, the temperature retrieval problem on the infrared region using satellite data provided by the High Resolution Infrared Sounder (HIRS) on board of NOAA series of weather satellites [2] is another example of a nonlinear retrieval problem. In this report, we will not enter on the details of either retrieval problem but will try to give an overview of possible methods for these problems.

The nonlinear retrieval discretization of the radiative transfer equation leads to the nonlinear model

$$\mathbf{T}_b = \mathbf{f}(\mathbf{x}) + \mathbf{e} \quad (37)$$

where \mathbf{x} is the quantity to be retrieved, and $\mathbf{f} : \mathcal{R}^n \rightarrow \mathcal{R}^m$ is a nonlinear vector function that maps the desired parameters to the measurement space. This problem also suffers from inherited ill-conditioning and is underdetermined since in most retrieval problems of interest $m < n$. This is further complicated by the fact that parameters of interest and measurements are related nonlinearly.

Nonlinear regularization methods are not as developed as regularization methods for linear problems. Methods for nonlinear regularization can in principle follow on ideas from linear regularization. The MMSE

(20) presented earlier is still the optimal estimator. One in principle can use the LMMSE (21) as it is actually done on the D-matrix algorithm [10] which is the operational algorithm for water vapor retrievals from SSM/T-2 measurements at AFWC. However, the LMMSE does not take into consideration the nonlinear features of the radiative transfer equation and as shown in [12, 13, 22] and iterative methods that account for the system model can improve over LMMSE.

6.1 Direct Nonlinear Regularization

Deterministic methods discussed earlier can be easily extended, in principle, to the nonlinear case [5, 21]. For instance Tikhonov regularization could be extended to the nonlinear case as follows. The Tikhonov regularized solution is the solution to the optimization problem.

$$\hat{\mathbf{x}}_\lambda = \arg \min_{\mathbf{x}} \left[\|\mathbf{T}_b - \mathbf{f}(\mathbf{x})\|^2 + \lambda^2 \|\mathbf{L}(\mathbf{x} - \mathbf{x}_o)\|^2 \right] \quad (38)$$

The regularized solution using the discrepancy principle will be

$$\hat{\mathbf{x}}_\alpha = \arg \min_{\mathbf{x}} \|\mathbf{L}(\mathbf{x} - \mathbf{x}_o)\|^2 \quad (39)$$

$$\text{Subject to } \|\mathbf{T}_b - \mathbf{f}(\mathbf{x})\|^2 \leq \alpha^2$$

6.2 Indirect Nonlinear Regularization

The methods discussed previously regularize the nonlinear problem directly. Another alternative in nonlinear least squares problems, is to linearize and then regularize the linearized problem. This is motivated by the use of the Gauss-Newton method [3] to solve nonlinear least squares problems where the search direction is computed by solving a linearization of the nonlinear problem.

The Gauss-Newton method for nonlinear least squares problems can be summarized as follows

Algorithm 6.1 Gauss-Newton Algorithm

- Given an initial guess \mathbf{x}_o and set $i = 0$. Iterate until convergence
 1. Compute the search direction \mathbf{s}_i by solving the linear least squares problem

$$\mathbf{s}_i = \arg \min_{\mathbf{s}} \|\mathbf{J}(\hat{\mathbf{x}}_i) \mathbf{s} - (\mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}_i))\|^2 \quad (40)$$

where $\hat{\mathbf{x}}_i$ is the estimate at the i -th iteration.

2. Update the estimate according to

$$\hat{\mathbf{x}}_{i+1} = \hat{\mathbf{x}}_i + \gamma_i \mathbf{s}_i \quad (41)$$

where γ_i is selected so that the cost function is sufficiently decreased.

3. If convergence criterion is satisfied stop, else $i = i + 1$ and go back to 1.

For ill-posed retrieval, problems the linear least squares problem (40) will also be ill-posed. This problem is linear and it is natural to apply regularization techniques for linear problems such as those discussed previously for the linear temperature retrieval problem. An application of this concept to temperature retrievals using IR data can be found in [19].

7 Conclusions and Final Comments

This paper presents some preliminary work in the application of regularization techniques to the linear problem of atmospheric temperature retrievals using microwave radiometry. The proposed techniques were evaluated using simulated data. The results obtained were quite encouraging. In particular, being able to estimate the location of the tropopause corner of the temperature profile even with bad initial guess and noisy data is a result not previously observed with other algorithms. Also, some approaches for the nonlinear retrieval problem were suggested.

Further validation with simulated and actual data is needed. Also other regularization algorithms could be studied to find out their performance in terms of location of the tropopause peak and accuracy of the estimates. In particular, all methods presented for the linear case were direct methods, iterative methods should be studied and their performance evaluated. Because of time limitation, we did not treated the issue of accuracy in any reasonable detail.

Nonlinear regularization methods should be studied in more detail. Our experience with current literature is that not enough attention have been paid to these methods and based on our experiences with the linear problem they could lead to better retrieval methods. We think this approach has the potential of being of benefit to water vapor retrievals using data from the SSM/T-2 sensor, temperature retrievals using IR data, and in unified retrievals. All of these can be fitted in the framework of nonlinear least squares.

Here the issue of computational efficiency was no addressed and certainly for use of any algorithm developed under this framework for operational retrieval it is of primary importance.

References

- [1] Bertsekas, D.P. *Nonlinear Programming*, Athena Scientific, 1995.
- [2] Cracknell, A.P. and L.W.B. Hayes. *Introduction to Remote Sensing*. Taylor & Francis Ltd., 1991.
- [3] Dennis, J.E. and R.B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice Hall, 1983.
- [4] Eldén, L. "Algorithms for regularization of ill-conditioned least-squares problems," *BIT*, Vol. 17, pp. 134-145, 1977.

- [5] Eriksson, J. *Optimization and Regularization of Nonlinear Least Squares Problems*. Ph.D. Thesis, Department of Computing Science, Umea University, Sweden, 1996.
- [6] Golub, G.H. and C.F. Van Loan. *Matrix Computations*, Second Edition, The John Hopkins University Press, 1989.
- [7] Groetsch, C.W. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman Publishing Inc., 1984.
- [8] Hansen, P.C. *Regularization Tools, a Matlab package for analysis and solution of discrete ill-posed problems; Version 2.0 for Matlab 4.0*, Technical University of Denmark, Report UNIC-92-03, March 1993.
- [9] Hansen, P.C. "Regularization Tools: A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems", *Numerical Algorithms*, Vol. 6, pp. 1-36, 1994.
- [10] Janssen, M.A. *Atmospheric Remote Sensing by Microwave Radiometry*, John Wiley & Sons, 1993.
- [11] Kazakos, D. and P. Papantoni-Kazakos. *Detection and Estimation*, Computer Science Press, 1990.
- [12] Al-Khalaf, A.K. *Retrieval of Atmospheric Water Vapor Profiles from the Special Sensor Microwave Temperature-2 (SSM/T-2)*. Ph.D. Thesis, Texas A & M University, May 1995.
- [13] Kuo, C.C. *Statistical Iterative Scheme for Estimating Atmospheric Relative Humidity Profiles from Microwave Radiometric Measurements*. Master Thesis, Massachusetts Institute of Technology, December 1988.
- [14] Lenoble, J. *Atmospheric Radiative Transfer*, A. Deepack Publishing, 1993.
- [15] Neu, T.J. "Defense meteorological satellite program microwave radiometer processing at air force global weather central," In *Proceedings of the First NMC/NESDIS/DOD Conference on DMSP Retrieval Products*, April 14-15, 1992. Phillips Laboratory, Report PL-TR-92-2191.
- [16] Paige, C.C. and M.A. Saunders. "Towards a generalized singular value decomposition," *SIAM Journal on Numerical Analysis*, Vol. 18, No. 3, 1981.
- [17] Rosenkranz, P.W., M.J. Komichak, and D.H. Staelin. "A Method for Estimation of Atmospheric Water Vapor Profiles by Microwave Radiometry." In *Journal of Applied Meteorology*, Vol. 21, pp. 1364-1370, 1982.
- [18] G.W. Stewart and J. Sun. *Matrix Perturbation Theory*. Academic Press, 1990.
- [19] Thompson, O.E. "Regularizing the satellite temperature-retrieval problem through singular-value decomposition of radiative transfer physics." In *Monthly Weather Review*, Vol. 20, pp. 2314-2328, 1992.

- [20] Tikhonov, A.N. and V.Y. Arsenin. *Solution of Ill-Posed Problems*, John Wiley & Sons, 1977.
- [21] Vogel, C.R. "An overview of numerical methods for nonlinear ill-posed problems" In *Inverse and Ill-posed Problems*, H.W. Engl and C.W. Groetsch, editors, Academic Press 1987.
- [22] Wilheit, T.T. "An algorithm for retrieving water vapor profiles in clear and cloudy atmospheres from 183 GHz radiometric measurements: simulation studies," In *Journal of Applied Meteorology*, Vol. 29, pp. 508-515, 1990.

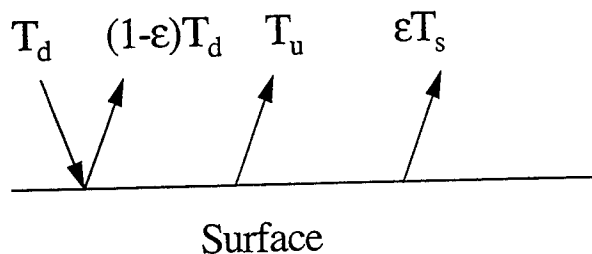


Figure 1: Boundary Condition for the Radiative Transfer Equation.

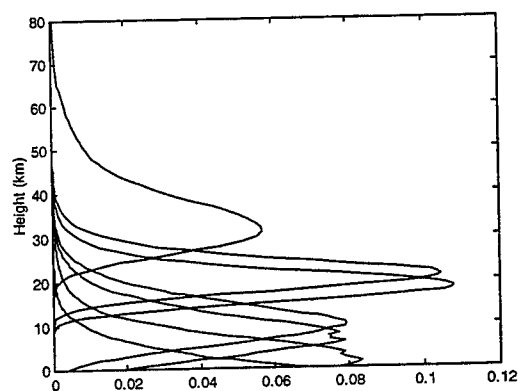


Figure 2: SSM/T-1 Weighting Functions.

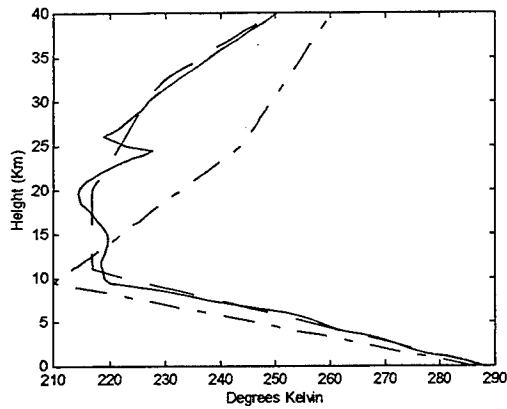


Figure 3: Tikhonov Regularization with 10 km Tropopause.

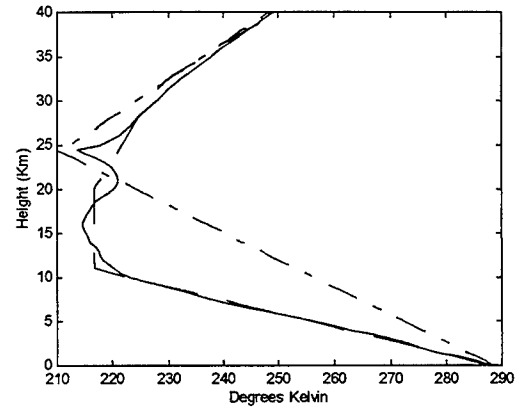


Figure 6: Tikhonov regularization with 25 km tropopause initial guess: noise free.

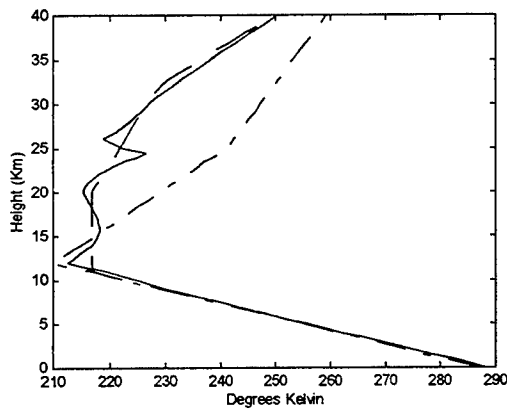


Figure 4: Tikhonov regularization with 12 km tropopause initial guess: noise free.

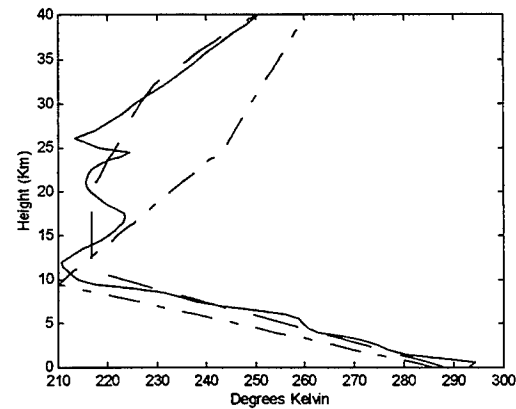


Figure 7: Tikhonov regularization with 10 km tropopause initial guess: noisy case.

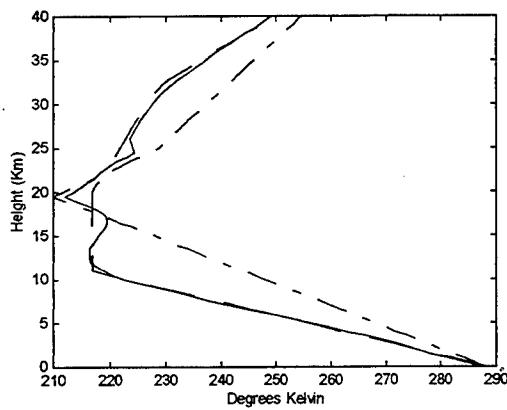


Figure 5: Tikhonov regularization with 20 km tropopause initial guess: noise free.

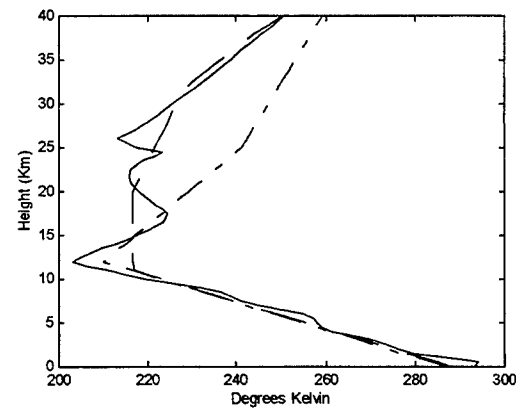


Figure 8: Tikhonov regularization with 12 km tropopause initial guess: noisy case.

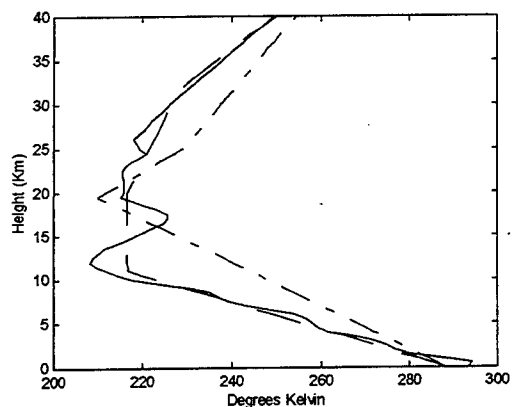


Figure 9: Tikhonov regularization with 20 km tropopause initial guess: noisy case.

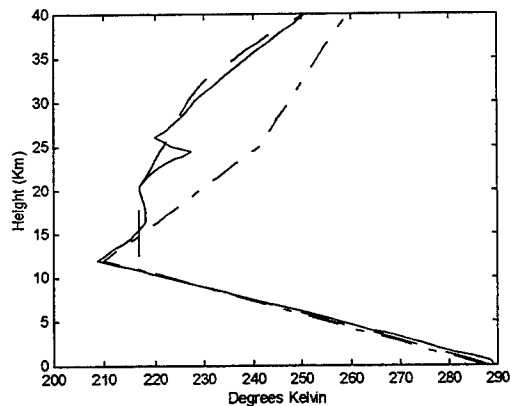


Figure 12: Discrepancy regularization with 12 km tropopause initial guess: noisy case.

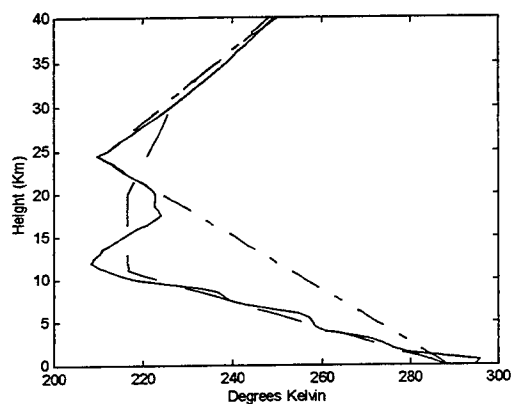


Figure 10: Tikhonov regularization with 25 km tropopause initial guess: noisy case.

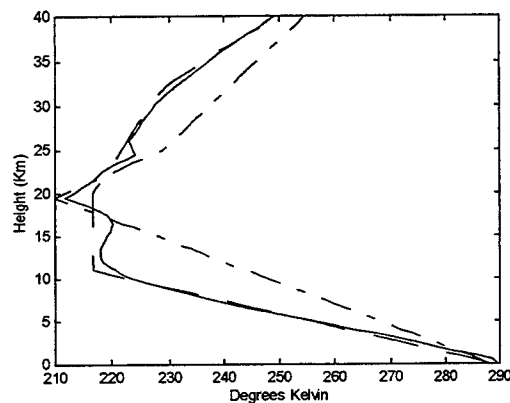


Figure 13: Discrepancy regularization with 20 km tropopause initial guess: noisy case.

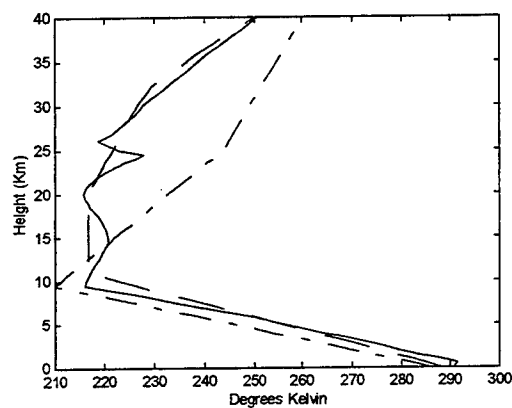


Figure 11: Discrepancy regularization with 10 km tropopause initial guess: noisy case.

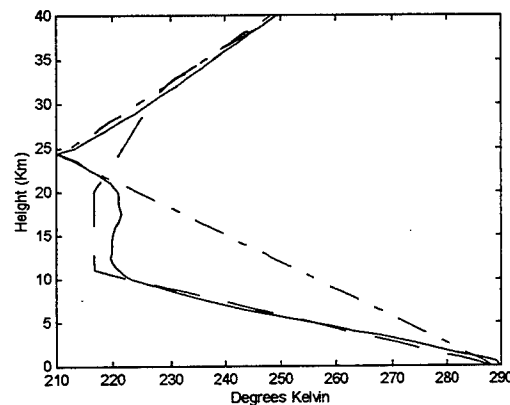


Figure 14: Discrepancy regularization with 25 km tropopause initial guess: noisy case.